Percolating, Cutting-down and Scaling Random Trees

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Gabriel Hernán Berzunza Ojeda

aus

Mexiko

Promotionskomitee

Prof. Dr.	Jean	Bertoin	(Vorsitz)
Prof. Dr.	Rémi	Abgrall	
Prof. Dr.	Valentin	Féray	
Prof. Dr.	Ashkan	Nikeghbali	
Prof. Dr.	Christina	Goldschmidt	(Gutachterin, University of Oxford)
Prof. Dr.	Gerónimo	Uribe Bravo	(Gutachter, Universidad Nacional Autónoma de México)

Zürich, 2017

Die vorliegende Arbeit wurde von der Mathematisch-naturwissenschaftlichen Fakultät der Universität Zürich im Herbstsemester 2016 als Dissertation angenommen.

Promotionskomitee: Prof. Dr. Jean Bertoin (Vorsitz), Prof. Dr. Rémy Abgrall, Prof. Dr. Valentin Féray, Prof. Dr. Ashkan Nikeghbali. Prof. Dr. Christina Goldschmidt (Gutachterin), Prof. Dr. Gerónimo Uribe Bravo (Gutachter).

A mis padres, Lourdes and Gabriel
A mis hermanos Alejandra y Adrian

"To me, you are still nothing more than a little boy who is just like a hundred thousand other little boys. And I have no need of you. And you, on your part, have no need of me. To you I am nothing more than a fox like a hundred thousand other foxes. But if you tame me, then we shall need each other. To me, you will be unique in all the world. To you, I shall be unique in all the world...."

— Antoine de Saint-Exupéry, The Little Prince

Summary

Trees are a fundamental notion in graph theory and combinatorics as well as a basic object for data structures, algorithms in computer science, statistical physics and the study of epidemics propagating in a network. In recent years, (random) trees have been the subject of many studies and various probabilistic techniques have been developed to describe their behaviors in different settings. In the first part of this thesis, we consider Bernoulli bond-percolation on large random trees. This means that each edge in the tree is removed with some fixed probability and independently of the other edges, inducing a partition of the set of vertices of the tree into connected clusters. We are interested in the supercritical regime, meaning informally that with high probability, there exists a giant cluster, that is of size comparable to that of the entire tree. We study the fluctuations of the size of such giant component, depending on the characteristics of the underlying tree, for two family of trees: b-ary recursive trees and scale-free trees. The approach relies on the analysis of the behavior of certain branching processes subject to rare neutral mutations.

In the second part, we study the procedure of cutting-down a tree. We destroy a large tree by removing its edges one after the other and in uniform random order until all the vertices are isolated. We then introduce a random combinatorial object, the so-called cut-tree, that represents the genealogy of the connected components created during the destruction. We investigate the geometry of this cut-tree, that depends of course on the nature of the underlying tree, and its implications on the multiple isolation of vertices. The study relies on the close relationship between the destruction process and Bernoulli bond percolation on trees.

In the last part of this thesis, we consider asymptotics of large multitype Galton–Watson trees. They are a natural generalization of the usual Galton-Watson trees and describe the genealogy of a population where individuals are differentiated by types that determine their offspring distribution. During the last years, research related to these trees has been developed in connection with important objects and models of growing relevance in modern probability such as random planar maps and non-crossing partitions, to mention just a few. We are more precisely interested in the asymptotic behavior of a function encoding these trees, the well-known height process. We consider offspring distributions that are critical and belong to the domain of attraction of a stable law. We show that these multitype trees behave asymptotically in a similar way as the monotype ones, and that after proper rescaling, they converge weakly to the same continuous random tree, the so-called stable Lévy tree. This extends the result obtained by Miermont [1] in the case of multitype Galton-Watson trees with finite variance.

Zusammenfassung

Bäume gehören zu den grundlegenden Begriffen der Graphentheorie und der Kombinatorik. Sie finden Anwendung in Datenstrukturen, Algorithmen in den Computerwissenschaften, statistischer Physik und bei der Untersuchung der Ausbreitung von Epidemien in Netzwerken. (Zufalls-)Bäume waren in den letzten Jahren immer wieder Gegenstand von Forschungen und verschiedenste Techniken wurden entwickelt um ihr Verhalten in unterschiedlichen Umgebungen zu beschreiben.

Im ersten Teil dieser Arbeit betrachten wir die Bernoulli Kantenperkolation auf grossen Zufallsbäumen. Das heisst, wir entfernen jede Kante des Baums mit einer fixen Wahrscheinlichkeit und unabhängig von den anderen Kanten. Auf diese Weise erhalten wir eine Partition der Menge der Knoten des Baumes in miteinander verbundene Cluster. Wir interessieren uns dabei für den superkritischen Verlauf, grob gesagt heisst das, dass mit einer hohen Wahrscheinlichkeit ein grosses Cluster existiert, dessen Grösse mit der Grösse des gesamten Baumes vergleichbar ist. Wir untersuchen die Schwankungen in der Grösse einer solchen Komponente in Abhängigkeit der Charakteristik des ihr zugrundeliegenden Baumes für zwei verschiedene Familie von Bäumen: b-näre rekursive Bäume und skalenfreie Bäume. Unser Ansatz beruht auf der Analyse des Verhaltens von Verzweigungsprozessen, wobei wir seltene, neutrale Mutationen zulassen.

Im zweiten Teil untersuchen wir ein Verfahren bei dem wir einen Baum fällen. Das heisst, wir zerstören einen grossen Baum indem wir nacheinander seine Kanten in einer zufälligen, gleichverteilten Reihenfolge entfernen bis alle seine Knoten isoliert sind. Wir führen den sogenannten cut-tree ein, ein zufälliges kombinatorisches Objekt, welcher die Abstammung der miteinander verbundenen Komponenten, die durch die Zerstörung entstanden sind beschreibt. Wir untersuchen die Geometrie dieses cut-trees, welche von dem ihr zugrundeliegenden Baumes und dessen Implikationen auf die Isolation der Knoten abhängt. Wir stützen uns dabei auf die enge Verbindung zwischen diesem Zestörungprozess und der Bernoulli Kantenperkolation.

Im letzten Teil der Arbeit untersuchen wir das asymptotische Verhalten von Multityp-Galton-Watson Bäumen. Diese sind die natürliche Verallgemeinerung des üblichen Galton-Watson Baums und beschreiben die Abstammung einer Population wenn die Individuen verschiedenen Typen angehören welche die Nachkommensverteilung bestimmen. In den letzten Jahren wurden Multityp-Bäume im Zusammenhang mit immer wichtiger werdenden Objekten und Modellen wie z.B. planaren Zufallskarten oder nicht-kreuzende Partitionen erforscht. Wir interessieren uns für das asymptotische Verhalten des Höheprozesses, einer Funktion welche relevante Informationen über den Baum enthält. Wir Betrachten kritische Nachkommensverteilungen welche zum Anziehungsbereich einer stabilen Verteilung gehören. Wir zeigen, dass diese Multityp-Bäume ein ähnliches asymptotisches Verhalten wie übliche Galton-Watson Bäume haben, und dass sie nach einer geeigneter reskalierung in Verteilung gegen den selben stetigen Zuffalsbaum, den sogenannten stablilen Lévy-Baum konvergieren. Dies ist eine Erweiterung eines Resultats von Miermont [1], welches den Fall von Multityp-Galton-Watson Bäume mit endlicher Varianz behandelt.

Acknowledgements

I would like to express my sincere appreciation to my advisor, Jean Bertoin, for the patient guidance, encouragement and advice he has provided throughout my time as his student. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly. Finally, I want to say that I admire not only his mathematical skills but also his human qualities. Muchas gracias, no encuentro las palabras para expresar todo mi agradecimiento.

I thank particularly Christina Goldschmidt and Gerónimo Uribe Bravo for reading and commenting this thesis. I would like to thank the members of my committee Ashkan Nikeghbali and Valentin Féray. I am also feel honored to have Rémi Abgrall, Grégory Miermont and Piere Nolin being part of my jury.

Gracias, Juan Carlos Pardo por haberme sugerido realizar el doctorado con Jean Bertoin. Sin duda alguna tu amistad y apoyo me ayudo mucho a lograr esta meta.

A mi familia que gracias a su apoyo pude concluir esta meta. Gracias a mis padres, Gabriel y Lourdes, por ayudarme a cumplir mis objetivos como persona y estudiante. A mi madre por hacer de mi una mejor persona a través de sus consejos, enseñanzas y amor. A mis hermanos, Alejandra y Adrian, por estar siempre presentes, acompañándome para poderme realizar. A Amado, por su amistad.

I am thankful to my friends from UZH and ETH for all the moments we shared inside and outside the academic life: Angelo, Alessandra, Alex, Avelio, Benjamin, Cyril, Erich, Igor, Julia, Marcel, Mayra, Rajat, Robin, Ron, Silvia and Xinji. Special thanks to my friend and officemate during these three years, Quan Shi.

I must also acknowledge and thank all the secretaries of the Institut für Mathematik which make life easier. I am especially grateful to Franziska Robmann for all her support since I arrived in Zürich.

A mis amigos Eduardo y Daniela por brindarme su confianza, amistad y apoyo durante todo este tiempo.

Finalmente, gracias a Lizeth por su compañía y por soportar todos los altibajos de este proyecto.

Contents

CHAPTER 1

Introduction

"Begin at the beginning," the King said gravely, "and go on till you come to the end: then stop."

— Lewis Carroll, Alice in Wonderland

This thesis consists of three main chapters whose common point is the study of different asymptotic behaviors for random trees. We first consider supercritical Bernoulli bond-percolation on large random trees, based on the article [2]. Then, we focus on another transformation of trees which bears close connection with percolation, the so-called destruction process; this is based on the work [3]. Finally, we study large multitype Galton-Watson trees, based on the article [4]. In this introductory chapter, we start by giving an informal description of the content of this thesis and provide some required background. The new results developed in this work are stated in the next three chapters which can be read independently.

1.1 Discrete random trees

Discrete random trees are special class of random graphs: they may be usually defined as connected graphs without cycles, and are often generated by some random procedure that describes how the edges or vertices are distributed. Let us explain briefly the terms appearing in this definition. A *graph* is a pair $G = (V, E)$, where V is the set of vertices and E is a subset of $V \times V$, called the edge set. An *edge* is a pair of different vertices. A *path* in a graph is a sequence of vertices where each successive pair of vertices is an edge in the graph. A finite path whose first and last vertices are the same is called *cycle*. Finally, we say that the graph is connected if there is a path from any of its vertices to any other. Therefore, most of the properties of trees are basic of graph theory whose foundation was laid down by Paul Erdős [5, 6, 7].

Let T be a tree with its respective set of vertices and edges. Throughout this work we consider rooted trees, that is, we distinguish a vertex which we call the root. This allows us to describe trees easily in terms of generations or levels. In words, the root is always the unique vertex at 0-th generation. The neighbors or children of the root constitute the first generation, and in general the vertices at distance k from the root form the k-th generation. The number of neighbors of a vertex v in T or the number of vertices that are adjacent to v is called the degree. Furthermore, vertices with degree one (except for the root) are usually called leaves or external vertices and the remaining ones internal vertices.

In the rest of this section we introduce two families of discrete random trees: increasing trees and Galton-Watson trees (more generally multitype Galton-Watson trees).

1.1.1 Increasing trees

A tree on an ordered set of vertices, say $[n] = \{1, \ldots, n\}$, is called increasing if when rooted at 1, the sequence of vertices along any branch from the root to a leaf increases; see Figure 1.1. The terminology comes from the fact that such trees can be constructed recursively, incorporating each vertex one after other in some order to built a growing tree; see for example Drmota [8] for details and further references.

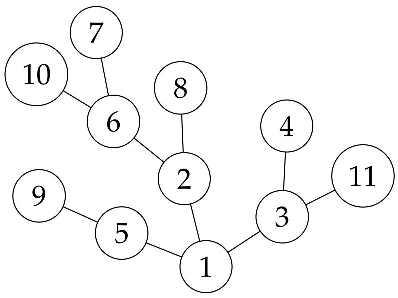

FIGURE 1.1: An example of increasing tree with set of vertices $[11] = \{1, \ldots, 11\}$.

Uniform random recursive tree. An important example, considered by Moon [9] and Flajolet and Steyaert [10], is the uniform random recursive tree. They arise for instance in computer science as data structures, or as simple epidemic models. An uniform random recursive tree on $[n]$ can be inductively constructed by the following algorithm: we start from the tree T_1 with a single vertex 1 (the root), and then successively, for every $k = 2, \ldots, n$, the vertex k is added to the tree T_{k-1}, attached by an edge to a vertex chosen uniformly at random among the $k-1$ already present.

Scale-free random trees. The scale-free random trees form a family of random trees that grows following a preferential attachment algorithm, and they are used commonly to model complex real-word networks. The name comes from Barabási and Albert [11] where this type of model was introduced in a more general version, in general yielding graphs rather than trees. It is important to point out that this model had been consider earlier by Szymańsky [12] under the name of nonuniform random recursive trees. Such trees are constructed recursively as follows: Fix a parameter $\beta \in (-1, \infty)$, and start for $n = 1$ from the unique tree T_1 on $\{1, 2\}$ which has a single edge connecting 1 and 2. Then suppose that T_n has been constructed for some $n \geq 1$, and for every $i \in \{1, \ldots, n+1\}$, denote by $d_n(i)$ the degree of the vertex i in T_n. Conditionally given T_n, the tree T_{n+1} is derived from T_n by incorporating an edge between the new vertex $n+2$ and a vertex $v_n \in \{1, \ldots, n+1\}$ chosen at random according to the law

$$\mathbb{P}(v_n = i | T_n) = \frac{d_n(i) + \beta}{2n + \beta(n+1)}, \qquad i \in \{1, \ldots, n+1\}.$$

The preceding expression defines a probability since the sum of the degrees of a tree with $n+1$ vertices equals $2n$. We observe that in the boundary case $\beta \to \infty$ where v_n becomes uniformly distributed on $[n+1]$, the algorithm yields a uniform random recursive tree. On the other hand, when $\beta = 0$, the procedure leads a so-called plane-oriented recursive tree; see for example [12].

b-ary recursive trees ($b \geq 2$). The b-ary recursive trees are mostly used for storing and searching data. The process to build a b-ary recursive tree starts at $n = 1$ from the tree T_1 with one internal vertex (which corresponds to the root) and b external vertices. Then, we suppose that T_n has been constructed for some $n \geq 1$ that is a tree with n internal vertices and $(b-1)n+1$ external ones (also called leaves). Then choose an external vertex uniformly at random and replace it by an internal vertex to which b new leaves are attached. In this way one continues. In the case $b = 2$, the algorithm yields a so-called binary search tree; see for instance Mahmoud [13] and Figure 1.2 for an ilustration.

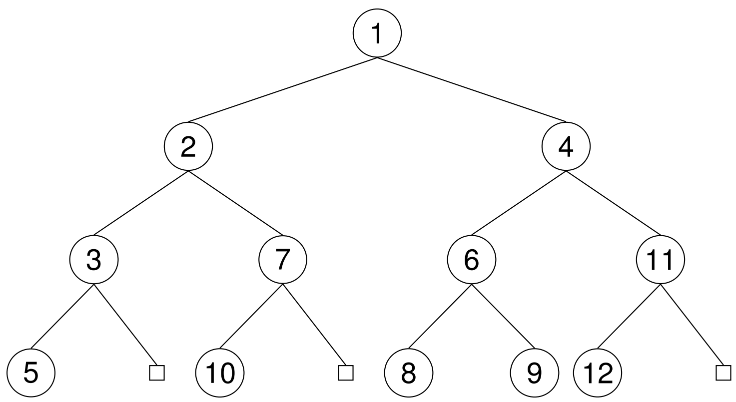

FIGURE 1.2: A binary search tree with set of vertices $[12] = \{1, 2, \ldots, 12\}$

1.1.2 Galton-Watson trees

Galton-Watson trees are an important and well-studied class of random trees in probability theory, they can be described as the genealogy of a Galton-Watson process started with one ancestor; these random trees are chosen to be rooted and ordered. They arises as building blocks of many different models of random graphs, such as Erdős and Rényi graphs or random planar maps, and appear also in combinatorics in relation to simply-generated trees. We next recall the standard formalism for family trees, first introduced by Neveu in [14].

Plane trees. Let U be the set of all labels:

$$U = \bigcup_{n=0}^{\infty} \mathbb{N}^n,$$

where $\mathbb{N} = \{1, 2, \ldots\}$ and with the convention $\mathbb{N}^0 = \{\varnothing\}$. An element of U is a sequence $u = u_1 \cdots u_j$ of positive integers, and we call $|u| = j$ the length of u (with the convention $|\varnothing| = 0$). If $u = u_1 \cdots u_j$ and $v = v_1 \cdots v_k$ belong to U, we write $uv = u_1 \cdots u_j v_1 \cdots v_k$ for the concatenation of u and v. In particular, note that $u\varnothing = \varnothing u = u$. For $u \in U$ and $A \subseteq U$, we let $uA = \{uv : v \in A\}$, and we say that u is a prefix (or ancestor) of v if $v \in uU$, in which case we write $u \vdash v$. Recall that the set U comes with a natural lexicographical order $\prec$, such that $u \prec v$ if and only if either $u \vdash v$, or $u = wu'$, $v = wv'$ with nonempty words u', v' such that $u'_1 < v'_1$.

A rooted planar tree $\mathbf{t}$ is a nonempty, finite subset of U which satisfies the following conditions:

1. $\varnothing \in \mathbf{t}$, we called it the root of $\mathbf{t}$.

2. For $u \in U$ and $i \in \mathbb{N}$, if $ui \in \mathbf{t}$ then $u \in \mathbf{t}$, and $uj \in \mathbf{t}$ for every $1 \leq j \leq i$.

We let $\mathbb{T}$ be the set of all rooted planar trees. We call vertices (or individuals) the elements of a tree $\mathbf{t} \in \mathbb{T}$, the length $|u|$ is called the height (generation) of $u \in \mathbf{t}$. We write $c_{\mathbf{t}}(u) = \max\{i \in \mathbb{Z}_+ : ui \in \mathbf{t}\}$ for the number of children of u. The vertices of $\mathbf{t}$ with no children are called leaves.

In addition to trees, we are also interested in forest. A forest $\mathbf{f}$ is a nonempty subset of U of the form

$$\mathbf{f} = \bigcup_k k\mathbf{t}_{(k)}, \qquad k \in \{1, 2, \dots, m\}, \;\; m \in \mathbb{N} \cup \{\infty\},$$

where $(\mathbf{t}_{(k)})$ is a finite or infinite sequence of trees, which are called the components of $\mathbf{f}$. In words, a forest may be thought of as a rooted tree where the vertices at height one are the roots of the forest components. We let $\mathbb{F}$ be the set of rooted planar forests. For $\mathbf{f} \in \mathbb{F}$, we define the subtree $\mathbf{f}_u = \{v \in U : uv \in \mathbf{f}\} \in \mathbb{T}$ if $u \in \mathbf{f}$, and $\mathbf{f}_u = \emptyset$ otherwise. We then observe that the tree components of $\mathbf{f}$ are $\mathbf{f}_1, \mathbf{f}_2, \dots$. We call $|u| - 1$ the height of $u \in \mathbf{f}$. Notice that that notion of height differs from the convention on trees because we want the roots of the forest components to be at height 0.

Galton-Watson trees. Let μ be a probability measure on $\mathbb{Z}_+$ such that $\mu(\{1\}) < 1$ (nondegenerating condition). We define the law $\mathbf{P}_\mu$ of a Galton-Watson tree with offspring distribution μ by

$$\mathbf{P}_\mu\left(T = \mathbf{t}\right) = \prod_{u \in \mathbf{t}} \mu\left(\{c_{\mathbf{t}}(u)\}\right),$$

where $T : \mathbb{T} \to \mathbb{T}$ is the identity map. This explicit formula is originally due to Otter [15]. We then say that a $\mathbb{T}$-value random variable T is a Galton-Watson tree with offspring distribution μ when it has law $\mathbf{P}_\mu$. Let

$$m = \sum_{z \in \mathbb{Z}_+} z\mu(\{z\})$$

be the mean number of the offspring distribution μ. Recall that μ is called sub-critical if $m < 1$, critical if $m = 1$ and super-critical if $m > 1$.

On the other hand, for an integer $r \geq 1$, we define $\mathbf{P}_{\mu,r}$ the law of Galton-Watson forest with r tree components and offspring distribution μ as the image measure of $\bigotimes_{j=1}^r \mathbf{P}_\mu$ by the map

$$(\mathbf{t}_{(1)}, \dots, \mathbf{t}_{(r)}) \longmapsto \cup_{k=1}^r k\mathbf{t}_{(k)},$$

i.e., it is the law that makes the identity map $F : \mathbb{F} \to \mathbb{F}$ the random forest whose trees components $F_1, \ldots, F_r$ are independent with law $\mathbf{P}_\mu$. Similarly, we can define the law of Galton-Watson forest with an infinite number of tree components.

We are further interested in multitype Galton-Watson trees, they are a natural generalization of Galton-Watson trees that describe the genealogy of a population where individuals are differentiated by types that determine their offspring distribution. The different types may correspond to actual different mutant forms of an organism or some other similar property. We then introduce the notion of d-type rooted planar tree by adding marks on the tree structure.

Multitype plane trees. For $d \in \mathbb{N}$, we call $[d] = \{1, \ldots, d\}$ the set of types. Then, a d-type rooted planar tree, or simply a multitype tree is a pair $(\mathbf{t}, e_\mathbf{t})$, where $\mathbf{t} \in \mathbb{T}$ and $e_\mathbf{t} : \mathbf{t} \to [d]$ is a function such that $e_\mathbf{t}(u)$ corresponds to the type of a vertex $u \in \mathbf{t}$. We let $\mathbb{T}^{(d)}$ be the set of d-type rooted planar trees. For $i \in [d]$, we write $c_\mathbf{t}^{(i)}(u) = \max\{j \in \mathbb{Z}_+ : uj \in \mathbf{t} \text{ and } e_\mathbf{t}(uj) = i\}$ for the number of offsprings of type i of $u \in \mathbf{t}$. Then, $c_\mathbf{t}(u) = \sum_{i \in [d]} c_\mathbf{t}^{(i)}(u)$ is the total number of children of $u \in \mathbf{t}$. Analogous definitions hold for d-type rooted planar forests $(\mathbf{f}, e_\mathbf{f})$, whose set will be denoted by $\mathbb{F}^{(d)}$.

Multitype Galton-Watson trees. For $d \in \mathbb{N}$, let $\boldsymbol{\mu} = (\mu^{(1)}, \ldots, \mu^{(d)})$ be a family of distributions on the space $\mathbb{Z}_+^d$ of integer-valued non-negative sequences of length d such that there exists at least one $i \in [d]$ so that

$$\mu^{(i)}\left(\left\{\mathbf{z} \in \mathbb{Z}_+^d : \sum_{j=1}^d z_j \neq 1\right\}\right) > 0$$

(nondegenerating condition). We define the law $\mathbf{P}_{\boldsymbol{\mu}}^{(i)}$ of a d-type Galton-Watson tree (or multitype Galton-Watson tree) rooted at a vertex of type $i \in [d]$ and with offspring distribution $\boldsymbol{\mu}$ by

$$\mathbf{P}_{\boldsymbol{\mu}}^{(i)}(T = \mathbf{t}) = \prod_{u \in \mathbf{t}} \frac{c_\mathbf{t}^{(1)}(u)! \ldots c_\mathbf{t}^{(d)}(u)!}{c_\mathbf{t}(u)!} \mu^{(e_\mathbf{t}(u))}\left(\left\{c_\mathbf{t}^{(d)}(u), \ldots, c_\mathbf{t}^{(d)}(u)\right\}\right),$$

where $T : \mathbb{T}^{(d)} \to \mathbb{T}^{(d)}$ is the identity map (see e.g., [16], or Miermont [1] for a formal construction of a probability measure on $\mathbb{T}^{(d)}$). We then say that a $\mathbb{T}^{(d)}$-value random variable T with law $\mathbf{P}_{\boldsymbol{\mu}}^{(i)}$ is a multitype Galton-Watson tree with offspring distribution $\boldsymbol{\mu}$ and root of type $i \in [d]$. Denote by

$$m_{ij} = \sum_{\mathbf{z} \in \mathbb{Z}_+^d} z_j \mu^{(i)}(\{\mathbf{z}\}), \quad \text{for} \quad i, j \in [d],$$

the mean number of children of type j, given by an individual of type i. We then let $\mathbf{M} := (m_{ij})_{i,j \in [d]}$ be the mean matrix of $\boldsymbol{\mu}$. Recall that if $\mathbf{M}$ is irreducible, then according to Perron-Frobenius theorem, $\mathbf{M}$ has a unique eigenvalue ρ which is simple, positive and with maximal modulus; see Chapter V of [17]. We then say that $\boldsymbol{\mu}$ is sub-critical if $\rho < 1$, critical $\rho = 1$ and super-critical if $\rho > 1$.

Similarly, for $\mathbf{x} = (x_1, \ldots, x_r)$ a finite sequence with terms in $[d]$, we define $\mathbf{P}_{\boldsymbol{\mu}}^{\mathbf{x}}$ the law of multitype Galton-Watson forest with roots of type $\mathbf{x}$ and with offspring distribution $\boldsymbol{\mu}$ as the image measure of

$\bigotimes_{j=1}^{r} \mathbf{P}_{\boldsymbol{\mu}}^{\mathbf{x}}$ by the map

$$(\mathbf{t}_{(1)}, \dots, \mathbf{t}_{(r)}) \longmapsto \cup_{k=1}^{r} k\mathbf{t}_{(k)},$$

i.e., it is the law that makes the identity map $F : \mathbb{F}^{(d)} \to \mathbb{F}^{(d)}$ the random forest whose trees components $F_1, \dots, F_r$ are independent with respective laws $\mathbf{P}_{\boldsymbol{\mu}}^{(x_1)}, \dots, \mathbf{P}_{\boldsymbol{\mu}}^{(x_d)}$. A similar definition holds for an infinite sequence $\mathbf{x} \in [d]^{\mathbb{N}}$.

1.2 Percolation on random trees

Consider a tree structure T_n on a finite set of vertices, say $[n] := \{1, \dots, n\}$, rooted at 1. We then perform Bernoulli bond percolation with parameter $p_n \in (0,1)$ that depends on the size of the original tree (i.e. the total number of vertices). So each edge is removed with probability $1 - p_n$ and independently of the other edges, inducing a partition of the set of vertices into connected clusters; see Figure 1.3. We are interested in the supercritical regime, in the sense that there exists a cluster whose size Γ_n satisfies that $n^{-1}\Gamma_n$ converges in law to some non-degenerate random variable, when $n \to \infty$. We then call Γ_n giant.

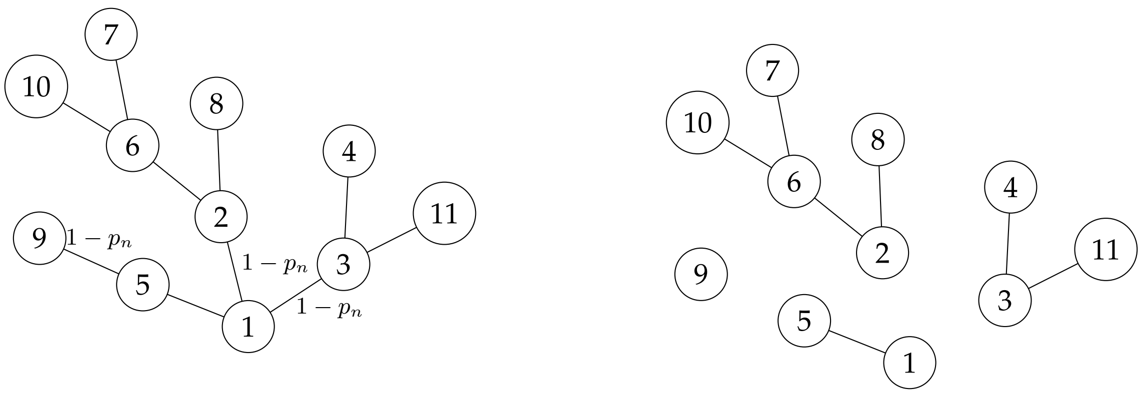

FIGURE 1.3: An increasing tree on the left, and the percolated tree on the right.

A result due to Bertoin [18] provides a simple criterion for the existence of giant percolation clusters in large trees, depending of course on the nature of the tree and regimes of percolation parameter. Informally, Bertoin has shown that for a fairly general family of trees, the supercritical regime corresponds to parameters of the form $p_n = 1 - c/\ell(n)$, where $c > 0$ is fixed and $\ell(n)$ is an estimate of the height of a typical vertex in the tree T_n. In several examples, the function ℓ is simply given by $\ell = \ln n$. For instance, this happens for some important families of random trees, such as uniform random recursive trees, binary search trees, scale-free random trees, etc.; see Section 1.1.1, Drmota [8] and Barabási [11]. Aldous [19] considered different class of examples, including the case when T_n is a Cayley tree of size n (i.e. a tree picked uniformly at random among the n^{n-2} trees on $[n]$), for which $\ell(n) = \sqrt{n}$.

Let us merely recall the case of uniform random recursive trees. Let $T_n^{(r)}$ be a uniform random recursive tree on $[n]$ and perform Bernoulli bond percolation with parameter

$$p_n = 1 - \frac{c}{\ln n}, \tag{1.1}$$

where $c > 0$ is fixed. It easy to show that this choice of the percolation parameter corresponds precisely to the supercritical regime, more precisely the size Γ_n of cluster containing the root satisfies that $\lim_{n\to\infty} n^{-1}\Gamma_n = e^{-c}$ in probability. We briefly sketch the proof of this result, referring to [18] for details. Pick a vertex u_n uniformly at random in $[n]$, and denote its distance to the root by h_n. It is well known that $h_n \sim \ln n$ (see Section 6.2.5 in [8]), and since the first moment of $n^{-1}\Gamma_n$ coincides with the probability that u_n is connected to the root, one gets

$$\mathbb{E}[n^{-1}\Gamma_n] = \mathbb{E}\left[\left(1 - \frac{c}{\ln n}\right)^{h_n}\right] \sim e^{-c}.$$

Similarly, let v_n be a second uniform vertex chosen independently of the first, then the easy fact that the height of the branch point of u_n and v_n remains stochastically bounded (see for instace [20]) yields the second moment estimate $\mathbb{E}[(n^{-1}\Gamma_n)^2] \sim e^{-2c}$, from which the law of large numbers for Γ_n follows. A natural problem is then to study its fluctuations.

Schweinsberg [21] (see also Bertoin [22] for an alternative approach) has shown that in the case of uniform random recursive trees, the fluctuations of the cluster containing the root are non-Gaussian. Specifically

$$(n^{-1}\Gamma_n - e^{-c})\ln n - ce^{-c}\ln\ln n \xrightarrow[n\to\infty]{d} -ce^{-c}(\mathcal{Z} + \ln c), \tag{1.2}$$

where the variable $\mathcal{Z}$ has the continuous Luria-Delbrück distribution, that is, its characteristic function is given by

$$\mathbb{E}[e^{i\theta\mathcal{Z}}] = \exp\left(-\frac{\pi}{2}|\theta| - i\theta\ln|\theta|\right), \qquad \theta \in \mathbb{R}.$$

We stress that this distribution arises in the limit theorem for sums of positive i.i.d. variables in the domain of attraction of completely asymmetric Cauchy process; see e.g. Möhle [23]. On the other hand, it was further observed in relation with a random algorithm for the isolation of the root, introduced by Meir and Moon [24] (in which edges are successively removed from the root cluster). In the context of uniform random recursive tree by Drmota et al. [25] and Iksanov and Möhle [26], and for random split trees by Holmgren [27, 28].

In Chapter 2, we investigate analogously the fluctuations of the giant cluster resulting from supercritical Bernoulli bond percolation in the case where T_n is a b-ary recursive trees with n internal vertices. More precisely, we perform Bernoulli bond percolation with parameter given by (1.1) that just as the case of the random recursive trees, it corresponds to the supercritical regime. This follows from the fact that the b-ary recursive trees have also logarithmic height, i.e. the height of typical vertex is approximately $\ell(n) = (b\ln n)/(b-1)$ (see Devroye [29]). Then, one can verify that percolation then produces a giant cluster whose size $C_0^{(p)}$ satisfies

$$\lim_{n\to\infty} n^{-1}C_0^{(p)} = e^{-\frac{b}{b-1}c} \qquad \text{in probability.}$$

We then establish the following limit theorem in law, which shows that the fluctuations of the giant cluster in the case of the b-ary recursive trees are also described by the continuous Luria-Delbrück distribution.

Theorem 1.1. *Set $\beta = b/(b-1)$, and assume that the percolation parameter p_n is given by (1.1). Then, there is the weak convergence*

$$(n^{-1}C_0^{(p)} - e^{-\beta c})\ln n - \beta c e^{-\beta c}\ln\ln n \xrightarrow[n\to\infty]{d} -\beta c e^{-\beta c}\mathcal{Z}_{c,\beta}$$

where

$$\mathcal{Z}_{c,\beta} = \mathcal{Z} - \kappa_\beta + \ln(\beta c) \tag{1.3}$$

with $\mathcal{Z}$ having the continuous Luria-Delbrück distribution,

$$\kappa_\beta = 1 - \frac{1}{\beta} + \frac{1}{\beta}\sum_{k=2}^{\infty}\frac{(\beta)_k}{k!}\frac{(-1)^k}{k-1}, \tag{1.4}$$

and $(x)_k = x(x-1)\cdots(x-k+1)$, for $k \in \mathbb{N}$ and $x \in \mathbb{R}$, is the Pochhammer function. In particular, for $b = 2$, i.e. for the binary search tree case, $\kappa_2 = 1$.

It should be noted the close similarity with the result for uniform recursive trees. It is remarkable that the normalizing functions and the limit in Theorem 1.1 only depend on the parameter $\beta = b/(b-1)$ through some constants. Observe that the left-hand side of (1.2) is the same as in Theorem 1.1 for $\beta = 1$; however the expressions (1.3) and (1.4) do not make sense $\beta = 1$!

The basic idea of Schweinsberg [21], for establishing the result (1.2) for uniform recursive trees relies on the estimation of the rate of decrease of the number of blocks in the Bolthausen-Sznitman coalescent, using the construction due to Goldschmidt and Martin [30] of the latter in terms of uniform recursive trees. On the other hand, the alternative approach of Bertoin [22] makes use on the remarkable coupling due to Iksanov and Möhle [26] connecting the Meir and Moon [24] algorithm for the isolation of the root, with a certain random walk in the domain of attraction of the completely asymmetric Cauchy process. These approaches depend crucially on the *splitting property* (see Section 3.1 in Bertoin [22]) which fails for the b-ary recursive trees. We thus have to use a different argument, although some guiding lines are similar to [22].

Essentially, we consider a continuous time version of the growth algorithm of the b-ary tree which bears close relations to Yule processes. The connection between recursive trees and branching processes is well-known, we make reference to Chauvin, et. al. [31] for the binary search trees and Bertoin and Uribe Bravo [32] for the case of scale-free trees. In this way, we adapt the recent strategy of [32]. Roughly speaking, incorporating percolation to the algorithm yields systems of branching processes with mutations, where a mutation event corresponds to disconnecting a leaf from its parent, and simultaneously replacing it by an internal vertex to which b new leaves are attached. Each percolation cluster size can then be thought of as a sub-population with some given genetic type. Hence the problem is reduced to study the fluctuations of the size of the sub-population with the ancestral type, which corresponds to the number of internal vertices connected to the root cluster.

This approach also allows us to study the fluctuations of the size of the giant cluster resulting from supercritical Bernoulli bond percolation on scale-free trees. More precisely, consider a scale-free tree $T_n^{(a)}$ of parameter $a \in (-1, \infty)$ with set of vertices $[n]$ (see Section 1.1.1). It has been observed by Bertoin and Uribe Bravo [32] that the percolation parameter (1.1) corresponds to the supercritical regime, and then the size of the cluster $\Gamma_n^{(\alpha)}$ containing the root satisfies

$$\lim_{n\to\infty} n^{-1}\Gamma_n^{(\alpha)} = e^{-\alpha c} \qquad \text{in probability},$$

where $\alpha = (1+a)/(2+a)$. We then have the following analogous result to Theorem 1.1.

Theorem 1.2. *Set $\alpha = (1+a)/(2+a)$, and assume that the percolation parameter p_n is given by (2.1). Then, there is the weak convergence*

$$\left(n^{-1}\Gamma_n^{(\alpha)} - e^{-\alpha c}\right)\ln n - \alpha c e^{-\alpha c}\ln\ln n \xrightarrow[n\to\infty]{d} -\alpha c e^{-\alpha c}\mathcal{Z}'_{c,\alpha}$$

where

$$\mathcal{Z}'_{c,\alpha} = \mathcal{Z} - \kappa'_\alpha + \ln(\alpha c)$$

with $\mathcal{Z}$ the continuous Luria-Delbrück distribution and

$$\kappa'_\alpha = 1 - \frac{1}{\alpha} + \frac{1}{\alpha}\sum_{k=2}^{\infty}\frac{(\alpha)_k}{k!}\frac{(-1)^k}{k-1}.$$

1.3 Cutting-down random trees

Let T_n be a tree on a finite set of vertices, say $[n] := \{1, \ldots, n\}$, rooted at 1. Imagine that we destroy it by cutting its edges one after the other, in a uniform random order. After $n-1$ steps, all edges have been destroyed and all the vertices are isolated. Meir and Moon [24, 33] initiated the study of such procedure by considering the number of cuts required to isolate the root, when the edges are removed from the current component containing this distinguished vertex. More precisely, they estimated the first and second moments of this quantity for two important trees families, Cayley trees and uniform random recursive trees. Concerning Cayley trees and other families of simply generated trees, a weak limit theorem for the number of cuts to isolate the root vertex was proven by Panholzer [34] and, in greater generality by Janson [35] who also obtained the result for complete binary trees [36]. Holmgren [27, 28] extended the approach of Janson to binary search trees and to the family of split trees. For random recursive trees a limit law was obtained, first by Drmota et al. [25] and reproved using a probabilistic approach by Iksanov and Möhle [26].

We observe that during the destruction process, the cut of an edge induces the partition of the subset (or block) that contains this edge into two sub-blocks of $[n]$. We then encode the destruction of T_n by a rooted binary tree, which we call the cut-tree and denote by $\text{Cut}(T_n)$. The cut-tree has internal vertices given by the non-singleton connected components which arise during the destruction, and leaves which correspond to the singletons $\{1\}, \ldots, \{n\}$ (these can be identified as the vertices of T_n). More precisely, the $\mathrm{Cut}(T_n)$ is rooted at the block $[n]$, then we build it inductively: we draw an edge between a parent block B and two children blocks B' and B'' whenever an edge is removed from the subtree of T_n with set of vertices B, producing two subtrees B' and B''. See Figure 1.4 for an illustration.

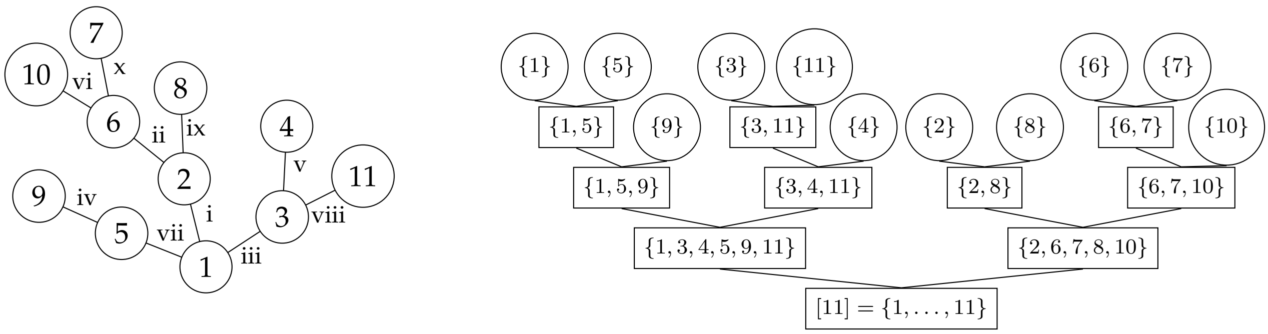

FIGURE 1.4: A tree of size eleven with the order of cuts on the left, and the corresponding cut-tree on the right

Roughly speaking, cut-trees describe the genealogy of connected components appearing in this edge-deletion process. They are especially useful in the study of the number of cuts needed to isolate any given subset of distinguished vertices, when the connected components which contain no distinguished points are discarded as soon as they appear. For instance, the number of cuts required to isolate k distinct vertices $v_1, \ldots, v_k$ coincides with the total length of the cut-tree reduced to its root and k leaves $\{v_1\}, \ldots, \{v_k\}$ minus $(k-1)$, where the length is measured as usual by the graph distance on $\mathrm{Cut}(T_n)$. This motivated the study of the cut-tree for several families of trees. Bertoin [37] considered the cut-tree of Cayley trees, and more generally, Bertoin and Miermont [38] dealt with critical Galton-Watson trees with finite variance and conditioned to have size n. Bertoin [39] studied the uniform random recursive trees, Dieuleveut [40] the Galton-Watson trees with offspring distribution belonging to the domain of attraction of a stable law of index $\alpha \in (1, 2]$, and Broutin and Wang [41] the so-called p-trees. Recently, Addario-Berry, Dieuleveut and Goldschmidt [42] developed a general framework for the study of cut-trees of real trees. They described the asymptotic behavior (in distribution) of the cut-trees when $n \to \infty$, for these classes of trees. However, we stress that the definition of the cut-tree differs slightly in these works, depending on the context.

In Chapter 3, we provide a general criterion for the convergence of the rescaled $\mathrm{Cut}(T_n)$ when the underlying tree T_n is star-shaped. Informally, we assume that the last common ancestor of two randomly chosen vertices is close to the root, after proper rescaling, with high probability. We consider also that T_n has a small height of order $o(\sqrt{n})$, in the sense that that the distance (the number of edges) between its root 1, and a typical vertex in T_n is of this order $o(\sqrt{n})$. In this direction, Sections 1.3.1 and 1.3.2 are devoted to recall the concept of real trees, and present the Gromov-Prokhorov topology in order to give a precise mathematical meaning to the convergence of rescaled discrete trees towards continuous objects. We then state the main result of Chapter 3 in Section 1.3.3.

1.3.1 Real trees

The following formal definition of real trees may be found for instance in Dress et. al. [43] or Evans [44]. A real tree is a metric space $(\mathcal{T}, d)$ which satisfies the following two properties:

1. For every $x, y \in \mathcal{T}$, there is a unique isometric map $\varphi_{x,y} : [0, d(x,y)] \to \mathcal{T}$ such that $\varphi_{x,y}(0) = x$ and $\varphi_{x,y}(d(x,y)) = y$.
2. For every $x, y \in \mathcal{T}$ and every continuous injective map $f : [0,1] \to \mathcal{T}$ such that $f(0) = x$ and $f(1) = y$, we have $f([0,1]) = \varphi_{x,y}([0, d(x,y)])$.

This is a continuous analog of the graph-theoretic definition of a tree as a connected graph with no cycle, that is, any two points in the metric space are linked by a geodesic path, which is the only simple path connecting these points, up to reparametrization. In this work, we consider rooted real tree $\mathcal{T} = (\mathcal{T}, d, \rho)$, i.e. a real tree $(\mathcal{T}, d)$ with a distinguished element $\rho \in \mathcal{T}$ called the root. In what follows, real trees will always be rooted, even if this is not mentioned explicitly. We also assume that the metric space $(\mathcal{T}, d)$ is complete and separable, unless we specify otherwise. In particular, when $(\mathcal{T}, d)$ is a compact metric space we call $\mathcal{T}$ compact real tree.

Example 1.1. *A construction of some particular cases of such metric spaces has been given by Aldous [19] and is described by Duquesne and Le Gall [45] in a more general setting. Let $g : [0,\infty) \to [0,\infty)$ be a continuous and compactly supported function such that $g(0) = 0$. We define a pseudo-distance on $[0,\infty)$ by setting*

$$d_g(s,t) = g(s) + g(t) - 2 \min_{r \in [s,t]} g(r)$$

for every $0 \le s \le t$ and the equivalence relation on $[0,\infty)$ by setting $s \overset{d_g}{\sim} t$ if and only if $d_g(s,t) = 0$. Consider the quotient space $\mathcal{T}_g = [0,\infty) \backslash \overset{d_g}{\sim}$ equipped with the distance induced by d_g; we keep the notation d_g for simplicity. Denote by $p_g : [0,\infty) \to \mathcal{T}_g$ the canonical projection. Then $\mathcal{T}_g = (\mathcal{T}_g, d_g, p_g(0))$ is a compact real tree (see Theorem 2.1 in [45]).

1.3.2 Gromov-Prokhorov topology

Let (E, δ) be a Polish space and denote by $C(E)$ the set of closed sets of E. Let $\mathcal{M}_1(E)$ denote the set of all Borel probability measures on E. We recall the definition of the Prokhorov metric: for every $\mu, \nu \in \mathcal{M}_1(E)$,

$$\delta_{\mathrm{P}}^E = \inf\{\varepsilon > 0 : \mu(A) \le \nu(A^\varepsilon) + \varepsilon \quad \text{and} \quad \nu(A) \le \mu(A^\varepsilon) + \varepsilon \quad \text{for any} \quad A \in C(E)\},$$

where $A^\varepsilon = \{x \in E : \delta(x, A) < \varepsilon\}$ is the ε-enlargement of A. It is well-known (see e.g. Billingsley [46]) that $(\mathcal{M}_1(E), \delta_{\mathrm{P}}^E)$ is a Polish space, and that the topology generated by δ_{P}^E is that of weak convergence.

We say that a quadruple (X, d, x, ν) is a pointed measured metric space (or sometimes pointed metric measure spaces) if (X, d) is a separable and complete metric space, $x \in X$ is a distinguished element called the root and ν is a Borel probability measure on (X, d). In order to compare two pointed measured metric spaces, we use the Gromov-Prokhorov distance: for every pointed measured metric spaces

(X, d, x, ν) and (X', d', x', ν'), we set:

$$d_{\text{GP}}(X, X') = \inf\{\delta_{\text{P}}^{E}(\phi \star \nu, \phi' \star \nu')\}$$

where the infimum is taken over all possible choices of a metric space (E, δ) and root-preserving isometric embeddings $\phi : X \to E$ and $\phi' : X' \to E$, and $\phi \star \nu$, $\phi' \star \nu'$ denote the the push-forward of μ, μ' by ϕ, ϕ', respectively.

Note that the function d_{GP} is a pseudo-distance; two pointed measured metric spaces (X, d, x, ν) and (X', d', x', ν') are called isometry-equivalent if there exists a root-preserving, bijective isometry that maps X onto X' and such that the push-forward of ν by the isometry is ν'. We denote by $\mathbb{M}$ the set of equivalence classes of pointed measured metric. It is also convenient to agree that for $a > 0$, aX denotes the same space (X, d, x, ν) but with distance rescaled by the factor a, i.e. (X, ad, x, ν). We recall also that the space $(\mathbb{M}, d_{\text{GP}})$ is a separable and complete metric space; see [47, 48].

Example 1.2. *A real tree $(\mathcal{T}, d, \rho)$ is a particular case of pointed metric space. Moreover, a real tree $(\mathcal{T}, \rho, d)$ equipped with a Borel probability measure ν on $(\mathcal{T}, d)$ is a pointed measured metric space which is known in the literature as a measured real tree. In particular, for the real tree $\mathcal{T}_g = (\mathcal{T}_g, d_g, p_g(0))$ described in Example 1.1, we consider the push-forward of the Lebesgue measure on the support of g by p_g.*

We then view the $\text{Cut}(T_n)$ for $n \geq 1$ as a sequence random variables with values in $\mathbb{M}$ (i.e. a sequence of real random tree). For convenience, we adopt a slightly different point of view for $\text{Cut}(T_n)$ than the usual for finite trees, focusing on leaves rather than internal nodes. Specifically, we set $[n]^0 = \{0, 1, \ldots, n\}$ where 0 corresponds to the root $[n]$ of $\text{Cut}(T_n)$ and $1, \ldots, n$ to the leaves (i.e. i is identified with the singleton $\{i\}$). We consider the random pointed metric measure space $([n]^0, \delta_n, 0, \mu_n)$ where δ_n is the random graph distance on $[n]^0$ induced by the cut-tree, 0 is the distinguished element, and μ_n is the uniform probability measure on $[n]$ extended by $\mu_n(0) = 0$. That is, μ_n is the uniform probability measure on the set of leaves of $\text{Cut}(T_n)$. We point out that the combinatorial structure of the cut-tree can be recovered from $([n]^0, \delta_n, 0, \mu_n)$, so by a slight abuse of notation, we refer to $\text{Cut}(T_n)$ as the latter pointed metric measured space.

Finally, let us give a simple characterization of the convergence in the sense of the pointed Gromov–Prokhorov topology due to Löhr [49]. A sequence (X_n, d_n, x_n, ν_n) of pointed measured metric spaces converges in the Gromov-Prokhorov sense to an element of $\mathbb{M}$, say $(X_\infty, d_\infty, x_\infty, \nu_\infty)$, if and only if the following holds: for $n \geq 0$, set $\xi_n(0) = \rho_n$ and let $\xi_n(1), \xi_n(2), \ldots$ be a sequence of i.i.d. random variables with law ν_n, then

$$(d_n(\xi_n(i), \xi_n(j)) : i, j \geq 0) \xrightarrow[n\to\infty]{d} (d_\infty(\xi_\infty(i), \xi_\infty(j)) : i, j \geq 0)$$

in the sense of finite-dimensional distributions and where $\xi_\infty(0) = \rho_\infty$ and $\xi_\infty(1)$, $\xi_\infty(2), \ldots$ is a sequence of i.i.d. random variables with law ν_∞.

One can interpret $(d_\infty(\xi_\infty(i), \xi_\infty(j)) : i, j \geq 0)$ as the matrix of mutual distances between the points of an i.i.d. sample of $(X_\infty, d_\infty, x_\infty, \nu_\infty)$. Moreover, it is important to point out that by the Gromov's reconstruction theorem in [50], the distribution of the above matrix of distances characterizes $(X_\infty, d_\infty, x_\infty, \nu_\infty)$ as an element of $\mathbb{M}$.

1.3.3 Convergence of the cut-tree

In this section, we state the main result of Chapter 3 which is a limit theorem for the rescaled $\mathrm{Cut}(T_n)$. First we need to introduce some notation. Recall that T_n is a tree with set of vertices $[n] = \{1, \dots, n\}$, rooted at 1. Let u, v be two independent uniformly distributed random vertices on $[n] = \{1, \dots, n\}$. Let d_n be the graph distance in T_n, and $\ell : \mathbb{N} \to \mathbb{R}_+$ be some function such that $\lim_{n\to\infty} \ell(n) = \infty$. We introduce the following hypothesis:

$$\frac{1}{\ell(n)}(d_n(1,u), d_n(u,v)) \xrightarrow[n\to\infty]{d} (\zeta_1, \zeta_1 + \zeta_2). \tag{H}$$

where ζ_1 and ζ_2 are i.i.d. variables in $\mathbb{R}_+$ with no atom at 0. This happens with ζ_i a positive constant for some important families of random trees, such as uniform recursive trees, scale-free random trees and b-ary recursive trees; see Drmota [8] and Barabási [11].

Remark 1.1. *We observe that*

$$d_n(u,v) = d_n(1,u) + d_n(1,v) - 2d_n(1, u \wedge v),$$

where $u \wedge v$ is the last common ancestor of u and v in T_n. Then, the condition (H) readily implies that

$$\lim_{n\to\infty} \ell(n)^{-1} d_n(1, u \wedge v) = 0,$$

in probability. Informally, the assumption (H) is saying that the branching points of T_n are close to the root (after rescaling), with high probability. We say that these trees are star-shaped.

We observe that (H) entails that

$$\frac{1}{\ell(n)} d_n(u,v) \xrightarrow[n\to\infty]{d} \zeta_1 + \zeta_2,$$

then we consider the next technical condition

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{\ell(n)}{d_n(u,v)} \mathbf{1}_{\{u \neq v\}}\right] = \mathbb{E}\left[\frac{1}{\zeta_1 + \zeta_2}\right] < \infty. \tag{H'}$$

We write

$$\lambda(t) = \mathbb{E}[e^{-t\zeta_1}], \qquad \text{for} \;\; t \geq 0,$$

for the Laplace transform of the random variable ζ_1 and denote $a = \mathbb{E}[1/\zeta_1]$ which can be infinite. We define the bijective mapping $\Lambda : [0, \infty) \to [0, a)$ by

$$\Lambda(t) = \int_0^t \lambda(s) \mathrm{d}s, \qquad \text{for} \;\; t \geq 0,$$

where $\Lambda(\infty) = \lim_{t\to\infty} \Lambda(t) = a$, and write Λ^{-1} for its inverse mapping.

Recall that we view $\mathrm{Cut}(T_n)$ as a pointed measured space.

Theorem 1.3. *Suppose that (H) and (H′) hold with ℓ such that $\ell(n) = o(\sqrt{n})$. Furthermore, assume that $a < \infty$. Then as $n \to \infty$, we have the following convergence in distribution in the sense of the pointed Gromov-Prokhorov topology:*

$$\frac{\ell(n)}{n}\mathrm{Cut}(T_n) \xrightarrow[n\to\infty]{d} I_\mu.$$

where I_μ is the pointed measure metric space given by the interval $[0, a]$, pointed at 0, equipped with the Euclidean distance, and the probability measure μ given by

$$\int_0^a f(x)\mu(\mathrm{d}x) = -\int_0^a f(x)\,\mathrm{d}\lambda \circ \Lambda^{-1}(x) \tag{1.5}$$

where f is a generic positive measurable function. The result still valid when $a = \infty$, and then one considers the interval $[0, \infty)$, pointed at 0, equipped with the same distance and measure.

Theorem 4.1 extends the result of Bertoin [39] for the convergence of the rescaled cut-tree associated with uniform random recursive trees in the sense of Gromov-Prokhorov topology. More precisely, it has been shown in [39] for a uniform random recursive tree $T_n^{(r)}$ of size n that upon rescaling the graph distance of $\mathrm{Cut}(T_n^{(r)})$ by a factor $n^{-1}\ln n$, the latter converges in probability in the sense of pointed Gromov-Hausdorff-Prokhorov distance to the unit interval $[0, 1]$ equipped with the Euclidean distance and the Lebesgue measure, and pointed at 0. The basic idea in [39] for establishing the result for uniform random recursive trees relies crucially on a coupling due to Iksanov and Möhle [26] that connects the destruction process in this family of trees with a remarkable random walk. However, this coupling is not fulfilled in general for the trees we are interested in, and thus we have to use a fairly different route.

On the other hand, we stress that Theorem 4.1 does not apply for the family of critical Galton-Watson trees conditioned to have size n considered by Bertoin and Miermont [38] and Dieuleveut [40] since they do not satisfy the condition (*H*), and the height of a typical vertex is not of the order $o(\sqrt{n})$. We believe that the threshold $\sqrt{n}$ appearing in this work is critical, and that for trees with larger heights (of order $\sqrt{n}$ or larger) the limit of their rescaled cut-tree is a random tree, and not a deterministic one. For instance, in the case when $T_n^{(c)}$ is a Cayley tree of size n, it has been shown in [37] that $n^{-1/2}\mathrm{Cut}(T_n^{(c)})$ converges in distribution to a Brownian Continuum Random tree, in the sense of Gromov-Hausdorff-Prokhorov. This uses crucially a general limit theorem due to Haas and Miermont [51] for so-called Markov branching trees. This has been extended in [38] to a large family of critical Galton-Watson trees with finite variance, and by Dieuleveut [40] when the offspring distribution belongs to the domain of attraction of a stable law of index $\alpha \in (1, 2]$, both in the sense of Gromov-Prokhorov. We point out that in [40] the limit is a stable random tree of index α.

Loosely speaking, our approach relies on the introduction of a continuous version of the cutting down procedure, where edges are equipped with i.i.d. exponential random variables and removed at a time given by the corresponding variable. Following Bertoin [52], we represent the destruction process up to a certain finite time as a Bernoulli bond-percolation, allowing us to relate the tree components with percolation clusters. We then develop the ideas in [52] used to analyze cluster sizes in supercritical percolation, and study the asymptotic behavior of the process that counts the number of edges which are remove from the root as time passed, which is closely related with the distance induced by the cut-tree. Finally, we make use of the convenient characterization of Löhr [49] for the Gromov-Prokhorov topology in order to get the convergence of the cut-tree.

In Chapter 3, we also present some applications of Theorem 1.3 on the isolation of multiple vertices, which extend the results of Kuba and Panholzer [53], and Baur and Bertoin [54] for uniform random recursive trees. More precisely, let $u_1, u_2, \dots$ denote a sequence of i.i.d. uniform random variables in $[n] = \{1, \dots, n\}$. We write $Z_{n,j}$ for the number of cuts which are needed to isolate $u_1, \dots, u_j$ in T_n.

Corollary 1.1. *Suppose that (H) and (H') hold with ℓ such that $\ell(n) = o(\sqrt{n})$. Then as $n \to \infty$, we have that*

$$\left(\frac{\ell(n)}{n} Z_{n,j} : j \geq 1\right) \xrightarrow[n\to\infty]{d} (\max(U_1, U_2, \dots, U_j) : j \geq 1)$$

in the sense of finite-dimensional distributions, where $U_1, U_2, \dots$ is a sequence of i.i.d. random variables with law μ given in (1.5).

In particular, when the hypotheses (H) and (H') hold with $\zeta_1 \equiv 1$, we observe that the variables $U_1, U_2, \dots$ have the uniform distribution on $[0, 1]$, and moreover, $\frac{\ell(n)}{n} Z_{n,j}$ converges in distribution to a beta$(j, 1)$ random variable. For example, this is the case of uniform random recursive trees, binary search trees, scale-free random trees, etc.; see Section 1.1.1.

As another application, for $j \geq 2$ we consider the algorithm for isolating the vertices $u_1, \dots, u_j$ with a slight modification, we discard the emerging tree components which contain at most one of these j vertices. We stop the algorithm when the j vertices are totally disconnected from each other, i.e. lie in j different tree components. We write $W_{n,2}$ for the number of steps of this algorithm until for the first time $u_1, \dots, u_j$ do not longer belong to the same tree component, moreover $W_{n,3}$ for the number of steps until the first time, the j vertices are spread out over three distinct tree components, and so on, up to $W_{n,j}$, the number of steps until the j vertices are totally disconnected.

Corollary 1.2. *Suppose that (H) and (H') hold with ℓ such that $\ell(n) = o(\sqrt{n})$. Then as $n \to \infty$, we have that*

$$\left(\frac{\ell(n)}{n} W_{n,2}, \dots, \frac{\ell(n)}{n} W_{n,j}\right) \xrightarrow[n\to\infty]{d} \left(U_{(1,j)}, \dots, U_{(j-1,j)}\right),$$

where $U_{(1,j)} \leq U_{(2,j)} \leq \cdots \leq U_{(j-1,j)}$ denote the first $j-1$ order statistics of an i.i.d. sequence $U_1, \dots, U_j$ of random variables with law μ given in (1.5).

As before, when (H) and (H') hold with $\zeta_1 \equiv 1$, the variables $U_1, U_2, \dots$ have the uniform distribution on $[0, 1]$, and then, $\frac{\ell(n)}{n} W_{n,j}$ converges in distribution to a beta$(1, j)$ random variable, and $\frac{\ell(n)}{n} W_{n,j}$ converges in distribution to a beta$(j-1, 2)$ law.

1.4 Scaling limits for multitype Galton-Watson trees

Galton-Watson branching processes are elementary models for the dynamics of a population, where at every generation each individual reproduces according to the same distribution, independently of the others. Natural generalizations of Galton-Watson processes are multitype Galton-Watson processes in which individuals have types that affect their reproduction law; see [55] for a introduction to these processes. An important object of study related to Galton-Watson processes is their underlying genealogical tree, the so-called Galton-Watson tree. Essentially, Galton-Watson trees describe the genealogical structure of the population as a planar rooted tree, with edges connecting parents to children; see Section 1.1.2 for a formal description. In a natural way, one defines a multitype Galton-Watson tree by incorporating the information about the types into the genealogical structure. Moreover, one can then consider multitype Galton-Watson forests, which are finite or infinite collections of genealogical trees related to independent multitype Galton-Watson processes.

It is well-known that the study of Galton-Watson trees is often reduced to analyze the behavior of renormalized functions that encode them. The idea of using encoding functions lies in the fact they have interesting probabilistic properties, and moreover there are many strategies available to prove a functional convergence. We describe in Section 1.4.1 three classical functions encoding the same rooted planar tree, namely the height process, the contour process, and the Lukasiewicz path. In Section 1.4.2, we recall two further Polish topologies on sets of trees. We then consider in Section 1.4.3 continuous analogous objects: we recall the definition of the Brownian continuum trees and the stable continuum tree. We present in Section 1.4.4 the results developed in Chapter 4.

1.4.1 Coding Galton-Watson trees

We fix for the whole section a rooted planar tree $\mathbf{t} \in \mathbb{T}$ and we denote by $\#\mathbf{t}$ its total progeny (or total number of vertices). Let us write by $\varnothing = u(0) \prec u(1) \prec \cdots \prec u(\#\mathbf{t}-1)$ the list of vertices of $\mathbf{t}$ in lexicographical order.

Height process. Recall that for a vertex $u \in \mathbf{t}$, we denote by $|u|$ its generation. We follows Le Gall and Le Jan [56] and define a process $H^{\mathbf{t}} = (H^{\mathbf{t}}_n, n \geq 0)$ by

$$H^{\mathbf{t}}_n = |u(n)| \qquad \text{for every } \ 0 \leq n < \#\mathbf{t},$$

with the convention that $H^{\mathbf{t}}_n = 0$ for $n \geq \#\mathbf{t}$. Sometimes it is convenient to think of $H^{\mathbf{t}}$ as the continuous function on $\mathbb{R}_+$, obtained by linear interpolation $t \to (1-\{t\})H^{\mathbf{t}}_{\lfloor t \rfloor} + \{t\}H^{\mathbf{t}}_{\lfloor t \rfloor+1}$, where $\{t\} = t - \lfloor t \rfloor$. The height process is thus the sequence of generations of the individuals of $\mathbf{t}$ visited in lexicographical order. It is easy to check that $H^{\mathbf{t}}$ fully characterizes the tree.

Contour process. To define the contour process of $\mathbf{t}$, imagine a particle that visits continuously all the edges at unit speed, from the left to the right, starting from the root (assuming that each edge has length one): after having reached $u(n)$, the particle goes to the individual $u(n+1)$, taking the shortest path. For $0 \leq t \leq 2(\#\mathbf{t}-1)$, we let $C^{\mathbf{t}}_t$ be the height of the particle at time t, i.e. its distance to the root. The process $C^{\mathbf{t}} = (C^{\mathbf{t}}_t, 0 \leq t \leq 2(\#\mathbf{t}-1))$ is called the contour process of $\mathbf{t}$. We observe that the contour process visits each edge of $\mathbf{t}$ exactly two times. Then, the contour process can be recovered from the height process through the following transform. We set

$$b_n = 2n - H^{\mathbf{t}}_n, \quad \text{for} \quad 0 \leq n < \#\mathbf{t} \quad \text{and} \quad b_{\#\mathbf{t}} = 2(\#\mathbf{t} - 1).$$

We observe that $0 = b_0 < b_1 < \cdots < b_{\#\mathbf{t}-1} < b_{\#\mathbf{t}} = 2(\#\mathbf{t} - 1)$. Then, for $0 \leq n < \#\mathbf{t} - 1$ and $t \in [b_n, b_{n-1})$,

$$C^{\mathbf{t}}_t = \begin{cases} H^{\mathbf{t}}_n - (t - b_n) & \text{if} \quad t \in [b_n, b_{n+1} - 1) \\ t - b_{n+1} + H^{\mathbf{t}}_{n+1} & \text{if} \quad t \in [b_{n+1} - 1, b_{n+1}), \end{cases}$$

and

$$C^{\mathbf{t}}_t = H^{\mathbf{t}}_{\#\mathbf{t}-1} - (t - b_{\#\mathbf{t}-1}) \qquad \text{if } t \in [b_{\#\mathbf{t}-1}, b_{\#\mathbf{t}}).$$

Lukasiewicz path. Define $W^{\mathbf{t}} = (W^{\mathbf{t}}_n, 0 \leq n < \#\mathbf{t})$ by $W^{\mathbf{t}}_0 = 0$ and for every $0 \leq n < \#\mathbf{t}$,

$$W^{\mathbf{t}}_{n+1} = W^{\mathbf{t}}_n + c_{\mathbf{t}}(u(n)) - 1,$$

where we recall that $c_{\mathbf{t}}(u)$ denotes the number of children of a vertex $u \in \mathbf{t}$. One can easily checks that $W^{\mathbf{t}}_n \geq 0$ for every $0 \leq n < \#\mathbf{t}$, and that $W^{\mathbf{t}}_{\#\mathbf{t}} = -1$. Observe that the jumps of $W^{\mathbf{t}}$ are not smaller than -1. On the other hand, the height process of a tree $\mathbf{t}$ can be recovered from its Lukasiewicz path by the following formula (see Corollary 2.2. of [56]):

$$H^{\mathbf{t}}_n = \text{Card}\left\{0 \leq j < n : W^{\mathbf{t}}_j = \inf_{j \leq k \leq n} W^{\mathbf{t}}_k\right\}, \qquad 0 \leq n < \#\mathbf{t}. \tag{1.6}$$

An important feature of this last encoding is that the Lukasiewicz path associated with a Galton-Watson tree with offspring distribution μ behaves as random walk starting from 0 with step distribution $\mu(\{\cdot + 1\})$ and stopped at the first hitting time of -1 (see e.g. Le Gall and Le Jan [56]). In contrast, the height process and the contour process are not Markovian in general.

Finally, we can extend the definition of the height process, the contour process and the Lukasiewicz path of a rooted planar forest $\mathbf{f} \in \mathbb{F}$; see for example [57]. Recall that for vertex $u \in \mathbf{f}$ of the forest, we call $|u| - 1$ its height or generation (the convention differs with the one of trees because we want the roots of $\mathbf{f}$ to be at height 0). Let $\mathbf{f}_1, \mathbf{f}_2, \dots$ be a forest the tree components of the forest $\mathbf{f}$. For $r \geq 1$, set $n_r = \#\mathbf{f}_1 + \cdots + \#\mathbf{f}_r$ with $n_0 = 0$. We define

$$H^{\mathbf{f}}_{n_r+k} = H^{\mathbf{f}_{r+1}}_k \qquad \text{and} \qquad W^{\mathbf{f}}_{n_r+k} = W^{\mathbf{f}_{r+1}}_k - r, \qquad 0 \leq k < \#\mathbf{f}_{r+1}$$

and

$$C^{\mathbf{f}}_{t+2n_r+2r} = C^{\mathbf{f}_{r+1}}_t, \qquad t \in [0, 2(\#\mathbf{f}_{r+1} - 1)).$$

Observe that $(n_r : r \geq 0)$ is the set of integers such that $H^{\mathbf{f}}_k = 0$. Consequently, the excursions of

$(H_n^{\mathbf{f}} : n \geq 0)$ above 0 are the $(H_{n_r+k}^{\mathbf{f}} : 0 \leq k \leq \#\mathbf{f}_{r+1})$. To the r-th tree of $\mathbf{f}$ corresponds the r-th excursion of above level zero of $H^{\mathbf{f}}$ and this excursion coincides with its height process.

We stress that similar definitions hold in a straightforward way for multitype plane tree and forest by forgetting their marks.

1.4.2 Gromov-Hausdorff-Prokhorov topology

Let (E, δ) be a Polish space and denote by $K(E)$ the set of compact sets of E. We next recall the Hausdorff distance on $K(E)$:

$$\delta_{\mathrm{H}}^{E} = \inf\{\varepsilon > 0 : A \subset B^{\varepsilon} \quad \text{and} \quad B \subset A^{\varepsilon}\} \quad \text{for every} \quad A, B \in K(E).$$

where $A^{\varepsilon} = \{x \in E : \delta(x, A) < \varepsilon\}$ is the ε-enlargement of A. We then use the Gromov-Hausdorff distance to compare two pointed compact metric spaces, say (X, d, x) and (X', d', x'): set

$$d_{\mathrm{GH}}(X, X') = \inf\{\delta_{\mathrm{H}}^{E}(\phi(X), \phi'(X')) \vee \delta_{\mathrm{H}}^{E}(\phi(x), \phi'(x'))\}$$

where the infimum is taken over all possible choices of a metric space (E, δ) and root-preserving isometric embeddings $\phi : X \to E$ and $\phi' : X' \to E$.

Recall that in Section 1.3.2, we introduce the Gromov-Prokhorov topology in order to compare two pointed metric measured spaces by equipping the spaces (X, d, x) and (X', d', x') with Borel probability measures ν and ν' respectively. We can consider both the metric and the measure with the Gromov–Hausdorff–Prokhorov distance to compare pointed compact measured metric spaces:

$$d_{\mathrm{GHP}}(X, X') = \inf\{\delta_{\mathrm{H}}^{E}(\phi(X), \phi'(X')) \vee \delta_{\mathrm{P}}^{E}(\phi \star \nu, \phi' \star \nu') \vee \delta_{\mathrm{H}}^{E}(\phi(x), \phi'(x'))\},$$

where again the infimum is taken over all possible choices of a metric space (E, δ) and root-preserving isometric embeddings $\phi : X \to E$ and $\phi' : X' \to E$, and $\phi \star \nu$, $\phi' \star \nu'$ denote the the push-forward of μ, μ' by ϕ, ϕ', respectively.

The functions d_{GH} and d_{GHP} as the Gromov-Prokhorov are also pseudo-distances. We say that two pointed compact metric spaces (X, d, x) and (X', d', x') are isometry-equivalent if there exists a root-preserving, bijective isometry that maps X onto X'. Recall also that two pointed compact measured metric spaces (X, d, x, ν) and (X', d', x', ν') are called isometry-equivalent if in addition the push-forward of ν by the isometry is ν'. We shall always implicitly identify two equivalent spaces: we denote by $\overline{\mathbb{M}}_{\mathrm{c}}$ the set of equivalence classes of pointed compact metric spaces and by $\mathbb{M}_c$ the set of equivalence classes of pointed compact measured metric spaces. Recall that the spaces $(\overline{\mathbb{M}}_{\mathrm{c}}, d_{\mathrm{GH}})$ and $(\mathbb{M}_{\mathrm{c}}, d_{\mathrm{GHP}})$ are separable and complete metric spaces; see [47, 48, 58].

Example 1.3. *A finite plane tree can be seen as a pointed (compact) metric space, when endowed with the graph distance. Furthermore, it can be turned into a real tree, by replacing each edge by a line segment of unit length; the associated function g is the linear interpolation of the contour process (see Section 3.2.1 in [51] for a detailed description).*

Denote next by $\overline{\mathbf{T}}_c$ the set of all equivalence classes of compact real trees, by $\mathbf{T}_c$ the set of all equivalence classes of compact measured real trees. It is well-know that the space $\overline{\mathbf{T}}_c$ is a closed subspace of $(\overline{\mathbb{M}}_c, d_{\mathrm{GH}})$ and the space $\mathbf{T}_c$ is a closed subspace of $(\mathbb{M}_c, d_{\mathrm{GHP}})$; see [59, 47].

Recall that in Example 1.1, we describe a procedure to construct real trees by using continuous and compactly supported functions. We conclude this subsection with an important result due to Duquesne and Le Gall that allows us to compare real trees $\mathcal{T}_g$ and $\mathcal{T}_{g'}$ coded by two different function g and g'.

Lemma 1.1. *([45]) Let $g, g' : [0,\infty) \to [0,\infty)$ be two continuous functions with compact support such that $g(0) = g'(0) = 0$. Then,*

$$d_{\mathrm{GH}}(\mathcal{T}_g, \mathcal{T}_{g'}) \leq 2\|g - g'\|, \tag{1.7}$$

where $\|\cdot\| = g \mapsto \sup_{x\in(0,\infty)} |g(x)|$ is the uniform norm.

Thanks to this result, we can indeed define a random real tree as the tree $\mathcal{T}_g$ coded by a random function g, i.e. that the map $g \mapsto \mathcal{T}_g$ is measurable. Moreover, we note that for a sequence of coding functions $(g_n)_{n\geq 1}$ such that $g_n \to g$ for the uniform topology, we have that $\mathcal{T}_{g_n} \to \mathcal{T}_g$ for the Gromov-Hausdorff-Prokhorov topology.

1.4.3 Continuum random trees

In this section, we recall the definition of Brownian continuum random tree and more generally the stable continuum random tree. We define them as real trees encoded by the analog in continuous-time of the discrete height process associated with a Galton-Watson tree.

The Brownian continuum tree. First, we recall the definition of the normalized Brownian excursion $\mathbb{e} = (\mathbb{e}_t, 0 \leq t \leq 1)$. Intuitively, the process $\mathbb{e}$ may be thought as a standard Brownian motion conditioned to remain nonnegative on $[0,1]$ and take the value 0 at time 1. However, the latter event has probability 0. On the other hand, there are several possible explicit descriptions of this process. For instance, let $B^{(3,1)}$ be a Bessel Bridge of dimension 3 over $[0,1]$, we then set

$$\mathbb{e}_t = B_t^{(3,1)}, \qquad \text{for } 0 \leq t \leq 1;$$

see for example Theorem 4.2 in [60]. The process $\mathbb{e}$ takes values in the Polish space $\mathbb{C}([0,1],\mathbb{R})$ of real-valued continuous functions on $[0,1]$ equipped with the uniform distance on all the compact subsets of $[0,1]$. We then define the Brownian continuum random tree (abbreviated as CRT for "Continuum Random Tree") as the random compact metric space $(\mathcal{T}_{2\mathbb{e}}, d_{2\mathbb{e}})$ coded by twice the normalized Brownian excursion. In the pioneer works [19, 61], Aldous introduced the continuum random tree as the limit of rescaled Galton-Watson trees conditioned on its total progeny for offspring distributions having finite variance. Specifically, he proved that their properly rescaled contour functions converge in distribution in the functional sense to the normalized Brownian excursion, which codes the continuum random tree as the contour function does for discrete trees.

Theorem 1.4. *([61]) Let μ be a critical probability measure on $\mathbb{Z}_+$ with variance $\sigma^2 \in (0,\infty)$. Let T be a Galton-Watson distributed according to $\mathbf{P}_\mu$. For every $n \geq 1$ for which $\mathbf{P}_\mu(\cdot|\#T = n)$ is well defined, we have* *that*

$$\left(\frac{\sigma}{\sqrt{2n}}C^T_{\#Tt}, 0 \leq t \leq 1\right) \xrightarrow[n\to\infty]{d} (2\mathbb{e}_t, 0 \leq t \leq 1)\,,$$

where the convergence is in distribution in $\mathbb{C}([0,1],\mathbb{R})$*, under* $\mathbf{P}_\mu(\cdot|\#T = n)$*.*

It then follows from Lemma 1.1 that, under the assumptions of Theorem 1.4, a Galton-Watson conditioned on its total progeny converges once its distances are properly rescaled to the CRT for the Gromov-Hausdorff-Prokhorov topology. More precisely, let T_n denote a Galton-Watson tree conditioned to have n vertices, for every $n \geq 1$ for which $\mathbf{P}_\mu(\#T = n) > 0$. Then,

$$\frac{\sigma}{\sqrt{n}}T_n \xrightarrow[n\to\infty]{d} \mathcal{T}_{2\mathbb{e}}$$

for the Gromov-Hausdorff-Prokhorov topology, where T_n is equipped with the graph distance and the uniform probability measure on the set of vertices.

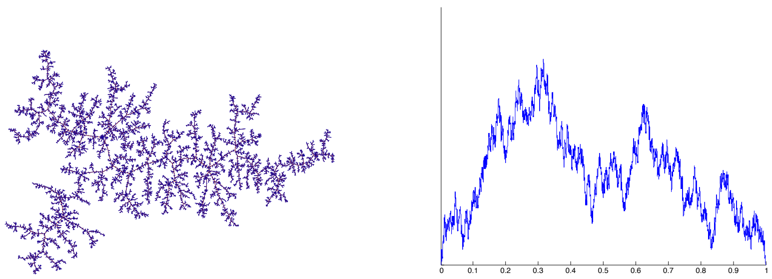

FIGURE 1.5: Simulation of $\mathcal{T}_{2\mathbb{e}}$ and $\mathbb{e}$ (by Igor Kortchemski).

This has motivated the study of the convergence of other rescaled paths obtained from Galton-Watson trees possibly with infinite variance, such as the Lukasiewicz path and the height process; see Section 1.4.1. Duquesne and Le Gall [62] obtained in full generality an unconditional version of Aldous' result. More precisely, they showed that the concatenation of rescaled height processes (or rescaled contour functions) converges in distribution to the so-called continuous-time height process associated to a spectrally positive Lévy process. We next recall the particular case when the offspring distribution belongs to the domain of attraction of a stable law of index $\alpha \in (1,2]$ which leads to the definition of the stable continuum random tree.

The stable continuum tree. Let μ be a probability measure on $\mathbb{Z}_+$. Recall that μ is in the domain of attraction of a stable law of index $\alpha \in (1,2]$ if: either the variance of μ is positive, finite and $\alpha = 2$, or $\mu([j,\infty)) = j^{-\alpha}L(j)$ for $j \geq 1$, where $L : \mathbb{R}_+ \to \mathbb{R}_+$ is a function such that $L(x) > 0$ for x large enough and $\lim_{x\to\infty} L(tx)/L(x) = 1$ for all $t > 0$ (such function is called slowly varying). In this framework, Duquesne [57] extended Theorem 1.4 by showing that the height processes and the Lukasiewicz path of Galton-Watson trees conditioned on having n vertices converge in distribution to the normalized excursion $Y^{\mathrm{exc}}_\alpha = (Y^{\mathrm{exc}}_t, 0 \leq t \leq 1)$ and continuous-time height process $H^{\mathrm{exc}}_\alpha = (H^{\mathrm{exc}}_t, 0 \leq t \leq 1)$ of a strictly stable spectrally positive Lévy process of index α; see Appendix A for a definition of these objects. Let $\mathbb{D}([0,1],\mathbb{R})$ the Polish space of càglàd functions on $[0,1]$ endowed with the Skorokhod topology.

Theorem 1.5. *([57]) Let μ be a critical probability measure on $\mathbb{Z}_+$ in the domain of attraction of a stable of index $\alpha \in (1, 2]$. Let T be a Galton-Watson tree distributed according to $\mathbf{P}_\mu$. For every $n \geq 1$ for which $\mathbf{P}_\mu(\cdot|\#T = n)$ is well defined, there exists a sequence of positive real numbers $B_n \to \infty$ such that*

$$\left(\left(\frac{1}{B_n} W^T_{\lfloor \#Tt \rfloor}, \frac{B_n}{n} H^T_{\#Tt}, \frac{B_n}{n} C^T_{\#Tt}\right), 0 \leq t \leq 1\right) \xrightarrow[n\to\infty]{d} \left(\left(Y^{\mathrm{exc}}_t, H^{\mathrm{exc}}_t, H^{\mathrm{exc}}_{t/2}\right), 0 \leq t \leq 1\right),$$

where the convergence is in distribution in $\mathbb{D}([0,1],\mathbb{R}) \otimes \mathbb{C}([0,1],\mathbb{R}) \otimes \mathbb{C}([0,1],\mathbb{R})$, under $\mathbf{P}_\mu(\cdot|\#T = n)$.

Following Duquesne and Le Gall [57] we define the α-stable continuum random tree as the random compact metric space $(\mathcal{T}_{H^{\mathrm{exc}}_\alpha}, d_{H^{\mathrm{exc}}_\alpha})$ encoded by the normalized excursion of the height process H^{exc}_α. We stress that in the Brownian case $H^{\mathrm{exc}}_2 = \sqrt{2} \cdot \mathbb{e}$, where $\mathbb{e}$ is the normalized Brownian excursion. Therefore, Theorem 1.5 could be also stated purely within the real tree formalism. Lemma 1.1 implies that under the above assumptions, a Galton-Watson conditioned on its total progeny converges after a proper rescaling to the α-stable continuum tree for the Gromov-Hausdorff-Prokhorov topology.

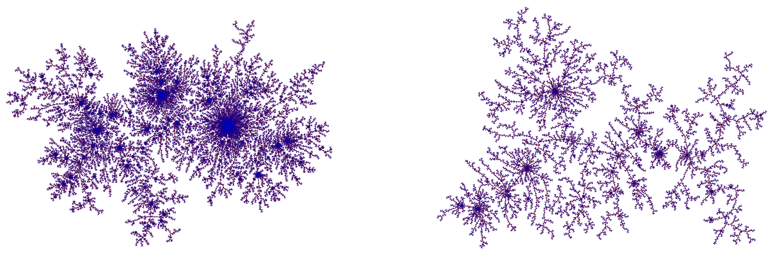

FIGURE 1.6: Simulation of $\mathcal{T}_{H^{\mathrm{exc}}_\alpha}$, respectively for α equals 1.1 and 1.5 (by Igor Kortchemski).

The stable continuum forest. Similarly, we can define the α-stable continuum random forest. First, we recall that a real forests is any countable collection of real trees; see Section 1.3.1. Then, the α-stable continuum random forest $\mathcal{F}_{H^{(\alpha)}}$ is the real forest coded by the continuous-time height process associated with a strictly α-stable spectrally positive Lévy process $H^{(\alpha)} = (H_t, t \geq 0)$; see Appendix A. Informally, following the idea of Example 1.1 every excursion above level zero of $H^{(\alpha)}$ corresponds to a tree component of $\mathcal{F}_{H^{(\alpha)}}$. In the Brownian case, $\alpha = 2$, the process $H^{(2)}$ is proportional to the reflected Brownian motion and such definition of a Brownian forest has already been introduced in Pitman [63], Proposition 7.8. The stable forest also arises as the limit of large Galton-Watson forests with offspring μ in the domain of attraction of a stable law of index $\alpha \in (1, 2]$. In this case, with slight abuse of notation, let us denote by $\mathbf{P}_\mu$ the law of a Galton-Watson forest with offspring distribution μ.

Theorem 1.6. *([62]) Let μ be a critical probability measure on $\mathbb{Z}_+$ in the domain of attraction of a stable of index $\alpha \in (1, 2]$. Let $F = (F_1, F_2, \dots)$ be a Galton-Watson forest distributed according to $\mathbf{P}_\mu$. Then, there exists a sequence of positive real numbers $B_n \to \infty$ such that*

$$\frac{B_n}{n}(F_1, \dots, F_n) \xrightarrow[n\to\infty]{d} \mathcal{F}_{H^{(\alpha)}}$$

for the Gromov-Hausdorff topology, where the Galton-Watson forest is equipped with the usual graph distance.

This is a consequence of Lemma 1.1, and Theorems 2.3.2 and 2.4.1 in [62] that show that the properly rescaled contour process and height process associated with a Galton-Watson forest, under the above assumptions, converge to the continuous-time process H_α which encodes the α-stable continuum random forest.

As a last remark, let us mention that in both theorems it is possible to give an explicit expression of the sequence $(B_n)_{n\geq 1}$ in terms of the offspring distribution μ. The formula is slightly complicated and we refer Theorem 1.10 in [64].

1.4.4 Convergence theorem for multitype Galton-Watson tree

Chapter 4 has been motivated by the following result of Miermont [1], which extends the classic results of monotype Galton-Watson trees. More precisely, Miermont establishes that the properly rescaled critical multitype Galton-Watson forest with finitely many types converges to the Brownian forest in the Gromov-Hausdorff sense, under the hypotheses that the offspring distribution is irreducible and has finite covariance matrix. Moreover, under an additional exponential moment assumption, he also established that conditionally on the number individuals of a given type, the multitype Galton-Watson tree converges to the CRT. More recently, de Raphelis [65] has extended Miermont's result for multitype Galton-Watson forest with infinitely many types, under similar assumptions. Informally speaking, these results claim that multitype Galton-Watson trees behave asymptotically in a similar way as the monotype ones, at least in the finite variance case. Therefore, this suggests that we should expect an analogous behavior for multitype Galton-Watson trees that satisfy different hypotheses.

In Chapter 4, we show an analogue result for critical multitype Galton-Watson forest with finitely many types whose offspring distribution is still irreducible, but belongs to the domain of attraction of a stable law. More precisely, for $d \in \mathbb{N}$, consider a d-type offspring distribution $\boldsymbol{\mu} = \left(\mu^{(1)}, \dots, \mu^{(d)}\right)$ on $\mathbb{Z}_+^d$ and define its Laplace transforms $\varphi = (\varphi^{(1)}, \dots, \varphi^{(d)})$ by

$$\varphi^{(i)}(\mathbf{s}) = \sum_{\mathbf{z}\in\mathbb{Z}_+^d} \mu^{(i)}(\{\mathbf{z}\}) \exp(-\langle \mathbf{z}, \mathbf{s}\rangle), \qquad \text{for } i \in [d],$$

where $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{R}_+^d$ and $\langle x, y\rangle$ is the usual scalar product of two vectors $x, y \in \mathbb{R}^d$. We let $\mathbf{0}$ be the vector of $\mathbb{R}_+^d$ with all components equal to 0, and for $i, j \in [d]$, we denote by

$$m_{ij} = -\frac{\partial \varphi^{(i)}}{\partial s_j}(\mathbf{0}) = \sum_{\mathbf{z}\in\mathbb{Z}_+^d} z_j \mu^{(i)}(\{\mathbf{z}\})$$

the mean number of children of type j, given by an individual of type i. Let $\mathbf{M} := (m_{ij})_{i,j\in[d]}$ be the mean matrix of $\boldsymbol{\mu}$, and $\mathbf{m}_i = (m_{i1}, \dots, m_{id}) \in \mathbb{R}_+^d$ be the mean vector of the measure $\mu^{(i)}$.

Recall also that if $\mathbf{M}$ is irreducible, then according to Perron-Frobenius theorem, $\mathbf{M}$ admits a unique eigenvalue ρ which is simple, positive and with maximal modulus; see Chapter V of [17]. We then say that $\boldsymbol{\mu}$ is sub-critical if $\rho < 1$, critical $\rho = 1$ and supercritical if $\rho > 1$.

Main assumptions. We assume that the offspring distribution $\boldsymbol{\mu} = \left(\mu^{(1)}, \dots, \mu^{(d)}\right)$ on $\mathbb{Z}_+^d$ satisfies the following conditions:

($\mathbf{H}_1$) $\boldsymbol{\mu}$ is irreducible, non-degenerate and critical. We then let $\mathbf{a} = (a_1, \dots, a_d)$ and $\mathbf{b} = (b_1, \dots, b_d)$ be the corresponding right and left 1-eigenvectors of $\mathbf{M}$ such that $\langle \mathbf{a}, 1 \rangle = \langle \mathbf{a}, \mathbf{b} \rangle = 1$.

($\mathbf{H}_2.\mathbf{1}$) Let Δ be a nonempty subset of $[d]$. For every $i \in \Delta$, there exists $\alpha_i \in (1, 2]$ such that the Laplace transform of $\mu^{(i)}$ satisfies

$$\psi^{(i)}(\mathbf{s}) := -\log \varphi^{(i)}(\mathbf{s}) = \langle \mathbf{m}_i, \mathbf{s} \rangle + |\mathbf{s}|^{\alpha_i} \Theta^{(i)}\left(\mathbf{s}/|\mathbf{s}|\right) + o(|\mathbf{s}|^{\alpha_i}), \qquad \text{as } |\mathbf{s}| \downarrow 0,$$

for $\mathbf{s} \in \mathbb{R}_+^d$ and where

$$\Theta^{(i)}(\mathbf{s}) = \int_{\mathbf{S}^d} |\langle \mathbf{s}, \mathbf{y} \rangle|^{\alpha_i} \lambda_i(\mathrm{d}\mathbf{y}),$$

with λ_i a finite Borel non-zero measure on $\mathbf{S}^d = \{\mathbf{y} \in \mathbb{R}^d : |\mathbf{y}| = 1\}$ such that for $\alpha_i \in (1, 2)$, λ_i has support in $\{\mathbf{y} \in \mathbb{R}_+^d : |\mathbf{y}| = 1\}$. We write $|\cdot|$ for the Euclidean norm.

($\mathbf{H}_2.\mathbf{2}$) For $i \in [d] \setminus \Delta$, the Laplace transform of $\mu^{(i)}$ satisfies

$$\psi^{(i)}(\mathbf{s}) := -\log \varphi^{(i)}(\mathbf{s}) = \langle \mathbf{m}_i, \mathbf{s} \rangle + o(|\mathbf{s}|^{\alpha_i}), \qquad \text{as } |\mathbf{s}| \downarrow 0.$$

where $\alpha_i = \min_{j \in \Delta} \alpha_j$.

Let us comment on these assumptions. Let $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots$ be a sequence of i.i.d. random variables on $\mathbb{Z}_+^d$ with common distribution $\mu^{(i)}$ satisfying ($\mathbf{H}_2.\mathbf{1}$). We observe that

$$-\log \mathbb{E}\left[\exp\left(-\left\langle \frac{1}{n^{1/\alpha_i}} \sum_{k=1}^{n} (\boldsymbol{\xi}_k - \mathbf{m}_i), \mathbf{s} \right\rangle\right)\right] \underset{n\to\infty}{\to} |\mathbf{s}|^{\alpha_i} \Theta^{(i)}\left(\mathbf{s}/|\mathbf{s}|\right), \qquad \mathbf{s} \in \mathbb{R}_+^d, \tag{1.8}$$

Then, we conclude that

$$\frac{1}{n^{1/\alpha_i}} \sum_{k=1}^{n} (\boldsymbol{\xi}_k - \mathbf{m}_i) \xrightarrow[n\to\infty]{d} \mathbf{Y}_{\alpha_i}, \tag{1.9}$$

where $\mathbf{Y}_{\alpha_i}$ is a α_i-stable random vector in $\mathbb{R}_+^d$ which Laplace exponent satisfies

$$\psi_{\mathbf{Y}_{\alpha_i}}(\mathbf{s}) = |\mathbf{s}|^{\alpha_i} \Theta^{(i)}\left(\mathbf{s}/|\mathbf{s}|\right), \qquad \mathbf{s} \in \mathbb{R}_+^d.$$

Sato's book [66] and [67] are good references for background on multivariate stable distributions. On the other hand, we notice from (1.8) that the equation (1.9) is equivalent to the hypothesis ($\mathbf{H}_2.\mathbf{1}$). We point out that in the monotype case, i.e. $d = 1$, the condition ($\mathbf{H}_2.\mathbf{1}$) may be thought as the analogous assumption made in [57] and [68], in order to get the convergence of the rescaled monotype Galton-Watson tree to the continuum stable tree. Finally, for $i \in [d] \setminus \Delta$, let $\mu^{(i)}$ be a measure that satisfies the hypothesis ($\mathbf{H}_2.\mathbf{2}$). We can rewrite the expression of its Laplace exponent in the following way

$$\psi^{(i)}(\mathbf{s}) := -\log \varphi^{(i)}(\mathbf{s}) = \langle \mathbf{m}_i, \mathbf{s} \rangle + |\mathbf{s}|^{\alpha_i} \Theta^{(i)}\left(\mathbf{s}/|\mathbf{s}|\right) + o(|\mathbf{s}|^{\alpha_i}), \qquad \text{as } |\mathbf{s}| \downarrow 0,$$

for $\mathbf{s} \in \mathbb{R}_+^d$ and where

$$\Theta^{(i)}(\mathbf{s}) = \int_{\mathbf{S}^d} |\langle \mathbf{s}, \mathbf{y}\rangle|^{\alpha_i} \lambda_i(\mathrm{d}\mathbf{y}),$$

with $\lambda_i \equiv 0$. Recall that $\alpha_i = \min_{j\in\Delta} \alpha_j$ for $i \in [d] \setminus \Delta$. This will be useful for the rest of the introduction.

Let us now write $\underline{\alpha} = \min_{i\in[d]} \alpha_i$ and $\bar{\lambda} = \sum_{i\in[d]} \mathbb{1}_{\{\underline{\alpha}=\alpha_i\}} a_i \lambda_i$. We then define

$$\bar{c} = (\langle \mathbf{a}, \mathbf{\Theta}(\mathbf{b})\rangle)^{1/\underline{\alpha}} = \left(\int_{\mathbf{S}^d} |\langle \mathbf{b}, \mathbf{y}\rangle|^{\underline{\alpha}} \bar{\lambda}(\mathrm{d}\mathbf{y})\right)^{1/\underline{\alpha}},$$

where $\mathbf{\Theta}(\mathbf{s}) = (\Theta^{(1)}(\mathbf{s})\mathbb{1}_{\{\underline{\alpha}=\alpha_1\}}, \dots, \Theta^{(d)}(\mathbf{s})\mathbb{1}_{\{\underline{\alpha}=\alpha_d\}}) \in \mathbb{R}_+^d$, for $\mathbf{s} \in \mathbb{R}_+^d$. We notice that $\bar{c} \not\equiv 0$ due to ($\mathbf{H}_2$.**1**). This constant will play a role similar to the constant defined in equation (2) of [1], i.e., it corresponds to the total variance of the offspring distribution $\boldsymbol{\mu}$, when the covariance matrices are finite.

We can now state the main result of Chapter 4. Recall from Section 1.4.3 that we write $\mathcal{F}_{H(\underline{\alpha})}$ for $\underline{\alpha}$-stable continuum random forest.

Theorem 1.7. *Let $F = (F_1, F_2, \dots)$ be a d-type GW forest distributed according to $\mathbf{P}^{\mathbf{x}}$, for some arbitrary $\mathbf{x} \in [d]^{\mathbb{N}}$. Then,*

$$\frac{\bar{c}}{n^{1-1/\underline{\alpha}}}(F_1, \dots, F_n) \xrightarrow[n\to\infty]{d} \mathcal{F}_{H(\underline{\alpha})}$$

for the Gromov-Hausdorff topology, where F is equipped with the graph distance.

Next, for $n \geq 0$, we let $\Upsilon_n^{\mathbf{f}}$ be the index of the tree component to which $u_{\mathbf{f}}(n)$ belongs.

Theorem 1.8. *For $i \in [d]$, let F be a d-type GW forest distributed according to $\mathbf{P}_{\boldsymbol{\mu}}^{\mathbf{i}}$, where $\mathbf{i} = (i, i, \dots)$. Then, under $\mathbf{P}_{\boldsymbol{\mu}}^{\mathbf{i}}$, we have the following convergence in distribution for the Skorokhod topology on the space $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ of càdlàd functions:*

$$\left(\frac{1}{n^{1/\underline{\alpha}}}\Upsilon^F_{\lfloor ns\rfloor}, s \geq 0\right) \xrightarrow[n\to\infty]{d} \left(-\frac{\bar{c}}{b_i} I_s, s \geq 0\right),$$

where I_s is the infimum at time s of the strictly $\underline{\alpha}$-stable spectrally positive Lévy process.

Let us briefly explain our approach, which relies on a remarkable decomposition of d-type forests into monotype forests introduced in [1]. The plan is to compare the corresponding height processes of the multitype Galton-Watson forest and the monotype Galton-Watson forest, and show that they are close for the Skorohod topology. We then pull back the known results of Duquesne and Le Gall [62] on the convergence of the rescaled height process of monotype Galton-Watson forests to the multitype Galton-Watson forest. Another important ingredient is the following estimate for the repartion of vertices of either type that may be of independent interest.

Proposition 1.1. *For $n \geq 0$, we let $\Lambda_i^{\mathbf{f}}(n)$ be the number of type i vertices standing before the $(n+1)$-th vertex in depth-first order. Then, for $i \in [d]$ and any $\mathbf{x} \in [d]^{\mathbb{N}}$, under $\mathbf{P}_{\boldsymbol{\mu}}^{\mathbf{x}}$, we have that*

$$\left(\frac{\Lambda_i^F(\lfloor ns\rfloor)}{n}, s \geq 0\right) \xrightarrow[n\to\infty]{} (a_i s, s \geq 0),$$

in probability, for the topology of uniform convergence over compact subsets of $\mathbb{R}_+$.

This result is know as convergence of types theorem, see for example [17].

We also present two applications of Theorem 1.7. We provide a limit theorem for the maximal height of a vertex in a multitype Galton-Watson tree that generalizes the result of Miermont [1] on the finite covariance case. The second application involves a particular multitype Galton-Watson tree, known as alternating two-type Galton-Watson tree, in which vertices of type 1 only give birth to vertices of type 2 and vice versa. We establish a conditioned version of Theorem 1.7 for this special tree. In particular, this family of two-type GW trees appears frequently in the study of random planar maps [69, 70], percolation on random triangulations and looptrees [71], non-crossing partitions [72], to mention couple of examples.

CHAPTER 2

Percolation on random trees

"It is an old maxim of mine that when you have excluded the impossible, whatever remains, however improbable, must be the truth."
— Sir Arthur Conan Doyle, The Adventure of the Beryl Coronet

In this chapter, we study the fluctuations of the giant cluster resulting from supercritical Bernoulli bond percolation on large b-ary trees and scale-free random trees. This is based on the article [2].

2.1 Introduction and main result

We start by recalling the percolation dynamics on trees. Consider a large tree T_n on a finite set of vertices say $[n] := \{1, \ldots, n\}$, rooted at 1. We then perform Bernoulli bond percolation with parameter $p_n \in (0, 1)$ that depends on the size of that tree (the total number of vertices), that is, each edge is removed with probability $1 - p_n$ and independently of the other edges, inducing a partition of the set of vertices into connected clusters. Recall that we are interested in the supercritical regime, in the sense that with high probability, there exists a giant cluster of size comparable to n, and its complement has also a size of order n. Theorem 1 in [18] shows that for fairly general families of trees,the supercritical regime corresponds to parameters of the form $p_n = 1 - c/\ell(n)$, where $c > 0$ a fixed parameter and $\ell(n)$ is an estimate of the height of a typical vertex in the tree structure.

Recall that in the case of the uniform random recursive trees $\ell(n) = \ln n$, so choosing the percolation parameter so that

$$p_n = 1 - \frac{c}{\ln n}, \tag{2.1}$$

where $c > 0$ is fixed, corresponds to the supercritical regime. Therefore, the size Γ_n of the largest cluster resulting from Bernoulli bond percolation satisfies $\lim_{n\to\infty} n^{-1}\Gamma_n = e^{-c}$ in probability. On the other hand, Schweinsberg [21] has shown that its fluctuations are non-Gaussian. More precisely,

$$(n^{-1}\Gamma_n - e^{-c}) \ln n - ce^{-c} \ln\ln n \Rightarrow -ce^{-c}(\mathcal{Z} + \ln c), \tag{2.2}$$

where $\Rightarrow$ means convergence in law as $n \to \infty$ and the variable $\mathcal{Z}$ has the continuous Luria-Delbrück distribution, i.e. its characteristic function is given by

$$\mathbb{E}[e^{i\theta\mathcal{Z}}] = \exp\left(-\frac{\pi}{2}|\theta| - i\theta \ln|\theta|\right), \qquad \theta \in \mathbb{R}.$$

The main purpose of this chapter is to investigate the case of large random b-ary recursive trees ($b \geq 2$). Recall that the process to build a b-ary recursive tree starts at $n = 1$ from the tree T_1 with one internal vertex (which corresponds to the root) and b external vertices. Then, we suppose that T_n has been constructed for some $n \geq 1$ that is a tree with n internal vertices and $(b-1)n+1$ external ones (also called leaves). Then choose an external vertex uniformly at random and replace it by an internal vertex to which b new leaves are attached. In this way one continues. Recall also that the case $b = 2$, the algorithm yields a so-called binary search tree. We consider that the size of the tree is the number of internal vertices.

Then we perform Bernoulli bond percolation with parameter given by (2.1) on a random b-ary recursive tree of size n, which corresponds precisely to the supercritical regime as the case of the random recursive trees. Roughly speaking, since the b-ary recursive trees have also logarithmic height, i.e. the height of typical vertex is approximately $\ell(n) = (b \ln n)/(b-1)$ (see Javanian and Vahidi-Asl [73]), one can verify that percolation then produces a giant cluster whose size $C_0^{(p)}$ (number of internal vertices) satisfies

$$\lim_{n\to\infty} n^{-1}C_0^{(p)} = e^{-\frac{b}{b-1}c} \qquad \text{in probability.}$$

We now state the central result of this chapter, which shows that the fluctuations of the giant cluster in the case of the b-ary recursive trees are also described by the continuous Luria-Delbrück distribution.

Theorem 2.1. *Set $\beta = b/(b-1)$, and assume that the percolation parameter p_n is given by (2.1). Then as $n \to \infty$, there is the weak convergence*

$$(n^{-1}C_0^{(p)} - e^{-\beta c}) \ln n - \beta c e^{-\beta c} \ln\ln n \Rightarrow -\beta c e^{-\beta c} \mathcal{Z}_{c,\beta}$$

where

$$\mathcal{Z}_{c,\beta} = \mathcal{Z} - \kappa_\beta + \ln(\beta c) \tag{2.3}$$

with $\mathcal{Z}$ having the continuous Luria-Delbrück distribution,

$$\kappa_\beta = 1 - \frac{1}{\beta} + \frac{1}{\beta}\sum_{k=2}^{\infty} \frac{(\beta)_k}{k!}\frac{(-1)^k}{k-1}, \tag{2.4}$$

and $(x)_k = x(x-1)\cdots(x-k+1)$, for $k \in \mathbb{N}$ and $x \in \mathbb{R}$, is the Pochhammer function. In particular, for $b = 2$, i.e. for the binary search tree case, $\kappa_2 = 1$.

The rest of this chapter is organized as follows. In Section 2.2, we consider a continuous time version of the growth algorithm of the b-ary tree which bears close relations to Yule processes. The connection between recursive trees and branching processes is well-known, we make reference to Chauvin, et.

al. [31] for the binary search trees and Bertoin and Uribe Bravo [32] for the case of scale-free trees. In this way, we adapt the recent strategy of [32]. Roughly speaking, incorporating percolation to the algorithm yields systems of branching processes with mutations, where a mutation event corresponds to disconnecting a leaf from its parent, and simultaneously replacing it by an internal vertex to which b new leaves are attached. Each percolation cluster size can then be thought of as a sub-population with some given genetic type. This lead us to investigate in the fluctuations of the size of the sub-population with the ancestral type, which corresponds to the number of internal vertices connected to the root cluster. Then in Section 2.3, we make the link with percolation on b-ary recursive trees in order to prove Theorem 2.1. Finally, we show in Section 2.4 that the present approach also applies to study the fluctuations of the size of the giant cluster for percolation on scale-free trees.

2.2 Yule process with rare mutations

The purpose of this section is to introduce a system of branching process with rare mutations, which is quite similar to the one considered in [32], although there are also some key differences (in particular, death may occur causing the extinction of sub-population with the ancestral type). Then we focus on estimating the size of the sub-population with the ancestral type, when the total population in the system grows and the mutation parameter depends of the size of the latter.

We consider a population in which each individual is either a clone (i.e. an individual with the ancestral type) or a mutant with some genetic type. A clone individual lives for an exponential time of parameter 1, and gives birth at its death to b clones with probability $p \in (0,1)$, or b mutants that share the same genetic type with probability $1-p$. A mutant individual lives for an exponential time of parameter 1, and gives birth at its death to b children of the same genetic of its parent. More precisely, the evolution of the population system is described by the process $\mathbf{Z}^{(p)} = (\mathbf{Z}^{(p)}(t) : t \geq 0)$, where

$$\mathbf{Z}^{(p)}(t) = (Z_0^{(p)}(t), Z_1(t), \dots), \qquad \text{for} \quad t \geq 0,$$

is a collection of nonnegative variables which represents the current size of the sub-populations. At the initial time, the sub-populations $Z_i(0)$ of type $i \geq 1$ are zero, and $Z_0^{(p)}(0) = b$ which is the size of the ancestral population. Formally, we take $\mathbf{Z}^{(p)}$ to be a pure-jump Markov chain whose transitions are described as follows. When at state $\mathbf{z} = (z_i : i \geq 0)$, our process jumps to a state $\tilde{\mathbf{z}} = (\tilde{z}_i : i \geq 0)$ where $\tilde{z}_j = z_j$ for $j \neq k$ and $\tilde{z}_k = z_k + (b-1)$ at rate

$$\begin{cases} pz_0 & \text{if} \quad k = 0, \\ z_k & \text{if} \quad k \neq 0. \end{cases}$$

This corresponds to a reproduction event in the sub-population with type k. Otherwise, the process jumps from $\mathbf{z}$ to $\hat{\mathbf{z}} = (\hat{z}_i : i \geq 0)$ at rate $(1-p)z_0$ where, if k is the first index such that $z_k = 0$, then $\hat{z}_0 = z_0 - 1$, $\hat{z}_k = b$, and $\hat{z}_j = z_j$ for $j \neq 0, k$. This corresponds to a mutation event of the sub-population with the ancestral type.

The process of the total size of the population in the system

$$Z(t) = Z_0^{(p)}(t) + \sum_{i\geq 1} Z_i(t), \qquad t \geq 0,$$

is distributed as a Yule process, where each individual lives for an exponential time of parameter 1 and gives birth at its death to b children, which then evolve independently of one another according to the same dynamics as their parent, no matter the choice of p. Clearly, the process of the size of the sub-population with the ancestral type $Z_0^{(p)}$ is a continuous time branching process, with reproduction law given by the distribution of $b\epsilon_p$, where ϵ_p stands for a Bernoulli random variable with parameter p. Moreover, if for $i \geq 1$, we write

$$a_i^{(p)} = \inf\{t \geq 0 : Z_i(t) > 0\},$$

for the birth time of the sub-population with type i, then each process

$$(Z_i(t - a_i^{(p)}) : t \geq a_i^{(p)})$$

is a branching process with the same reproduction law as Z starting from b. Indeed, the different populations present in the system (i.e., those with strictly positive sizes) evolve independently of one another. The following statement is just a formal formulation of the previous observation which should be plain from the construction of $\mathbf{Z}^{(p)}$; it is essentially Lemma 1 in [32].

Lemma 2.1. *The processes $(Z_i(t - a_i^{(p)}) : t \geq a_i^{(p)})$ for $i \geq 1$ form a sequence of i.i.d. branching process with the same law as Z and with starting value b. Further, this sequence is independent of that of the birth-times $(a_i^{(p)})_{i\geq 1}$ and the process $Z_0^{(p)}$ of the sub-population with ancestral type.*

We are now ready to present the main result of this section. We henceforth assume that the parameter $p = p_n$ is given by (2.1) and for simplicity, we write p rather than p_n, omitting the integer n from the notation. We consider the time

$$\tau(n) = \inf\{t \geq 0 : Z(t) = (b-1)n + 1\},$$

when the total population has size $(b-1)n + 1$. The size of the sub-population with the ancestral type at this time is given by

$$G_n := Z_0^{(p)}(\tau(n)).$$

Theorem 2.2. *Set $\beta = b/(b-1)$. As $n \to \infty$, there is the weak convergence*

$$\left(n^{-1}G_n - \frac{1}{\beta - 1}e^{-\beta c}\right)\ln n - \frac{\beta}{\beta - 1}ce^{-\beta c}\ln\ln n \Rightarrow -\frac{\beta}{\beta - 1}ce^{-\beta c}\left(\mathcal{Z}_{c,\beta} + 1 - \frac{1}{\beta}\right),$$

where $\mathcal{Z}_{c,\beta}$ is the random variable defined in (2.3).

We stress that this result also allows us to deduce the fluctuations of the number of mutants in the total population, since this quantity is given by $(b-1)n + 1 - G_n$.

The rest of this section is devoted to the proof of Theorem 2.2. Our approach is similar to that in [22]. Broadly speaking, we divide the study of the fluctuations in two well-defined phases. The crucial point is to obtain a precise estimate of the number Δ_n of mutants when the total population of the system attains the size $(b-1)\lfloor \ln^4 n \rfloor + 1$; this can be viewed as the germ of the fluctuations of $(b-1)n + 1 - G_n$. Then, we resume the growth of the system from size $(b-1)\lfloor \ln^4 n \rfloor + 1$ to the size $(b-1)n + 1$ and observe that the sub-population with the ancestral type grows essentially regularly. We point out that even though the study of these two phases plays a key role in [22], the tools developed here to deal with each phase are much different from those used there.

2.2.1 The germ of fluctuations

In this first phase, we observe the growth of the system of branching processes until the time

$$\tau(\ln^4 n) = \inf\{t \geq 0 : Z(t) = (b-1)\lfloor \ln^4 n \rfloor + 1\},$$

which is when the total size of the population reaches $(b-1)\lfloor \ln^4 n \rfloor + 1$, and our purpose in this section is to estimate precisely the number Δ_n of mutants in the total population at this time, that is

$$\Delta_n = (b-1)\lfloor \ln^4 n \rfloor + 1 - Z_0^{(p)}(\tau(\ln^4 n)).$$

We stress that the threshold $(b-1)\ln^4 n + 1$ is somewhat arbitrary, and any power close to 4 of $\ln n$ would work just as well. However, as is remarked by Bertoin in [22], it is crucial to choose a threshold which is both sufficiently high so that fluctuations are already visible, and sufficiently low so that one can estimate the germ with the desired accuracy.

We start by setting down the key results that lead us to the main result of this section, in order to give an easier articulation of the argument. In this direction, it is convenient to introduce the number $\Delta_{0,n}$ of mutants at time

$$\tau_0(\ln^4 n) = \inf\{t \geq 0 : Z_0^{(p)}(t) = (b-1)\lfloor \ln^4 n \rfloor + 1\},$$

which is when the size of the sub-population with the ancestral type reaches $(b-1)\lfloor \ln^4 n \rfloor + 1$, i.e.

$$\Delta_{0,n} = Z(\tau_0(\ln^4 n)) - (b-1)\lfloor \ln^4 n \rfloor - 1.$$

This will be useful since the distribution of $\Delta_{0,n}$ is easier to estimate than that of Δ_n. Then, we establish the following limit theorem in law that relates the fluctuations of $\Delta_{0,n}$ with the continuous Luria-Delbrück variable $\mathcal{Z}$.

Proposition 2.1. *As $n \to \infty$, there is the weak convergence*

$$\frac{\Delta_{0,n}}{\ln^3 n} - 3\frac{\beta}{\beta-1} c \ln\ln n \Rightarrow \frac{\beta}{\beta-1} c \left(\mathcal{Z}_{c,\beta} + 1 - \frac{1}{\beta} \right)$$

where $\mathcal{Z}_{c,\beta}$ is the random variable defined in (2.3).

As we are interested in estimate the number Δ_n of mutants in the total population at time $\tau(\ln^4 n)$, and we know the behavior of $\Delta_{0,n}$, the purpose of the next lemma is to point out that these two quantities are close enough when $n \to \infty$. We need to introduce the notation:

$$A_n = B_n + o(f(n)) \quad \text{in probability},$$

where A_n and B_n are two sequences of random variables and $f : \mathbb{N} \to (0, \infty)$ is a function, to indicate that $|A_n - B_n|/f(n) \to 0$ in probability when $n \to \infty$.

Lemma 2.2. *We have*

$$\Delta_n = \Delta_{0,n} + o(\ln^3 n) \quad \textit{in probability}.$$

It then follows from Proposition 2.1 that Δ_n and $\Delta_{0,n}$ have the same asymptotic behavior. Specifically:

Corollary 2.1. *As $n \to \infty$, there is the weak convergence*

$$\frac{\Delta_n}{\ln^3 n} - 3\frac{\beta}{\beta - 1} c \ln \ln n \Rightarrow \frac{\beta}{\beta - 1} c \left(\mathcal{Z}_{c,\beta} + 1 - \frac{1}{\beta} \right)$$

where $\mathcal{Z}_{c,\beta}$ is the random variable defined in (2.3).

The above result will be sufficient for our purpose. We now prepare the ground for the proofs of Proposition 2.1 and Lemma 2.2. Recall that we wish to study the behavior of the number $\Delta_{0,n}$ of mutants at time $\tau_0(\ln^4 n)$, which is easier than that of Δ_n, thanks to Lemma 2.1. In words, at time $\tau_0(\ln^4 n)$ there is an independence property between the mutant sub-populations, and the process that counts the number of mutation events, which allows us to express $\Delta_{0,n}$ as a random sum of independent Yule processes. Clearly, the above is not possible at time $\tau(\ln^4 n)$ due to the lack of independence within the sub-populations. Formally, we start by writing

$$M(t) = \max\{i \geq 1 : Z_i(t) > 0\}$$

for the number of mutations that have occurred before time $t \geq 0$. Lemma 2.1 ensures that M is independent of the processes $(Z_i(t - a_i^{(p)}) : t \geq a_i^{(p)})$ for $i \geq 1$. In addition, we note that the jump times of M are in fact $a_1^{(p)} < a_2^{(p)} < \cdots$. This enables us to express the total mutant population at time t as,

$$Z_{\mathrm{m}}(t) = \sum_{i=1}^{M(t)} Z_i(t - a_i^{(p)}),$$

and we are thus interested in

$$\Delta_{0,n} = Z_{\mathrm{m}}(\tau_0(\ln^4 n)). \tag{2.5}$$

We now turn our attention to study the fluctuations of $\Delta_{0,n}$ through the analysis of its characteristic function. In this direction, we will be mainly interested in the following feature of $Z_m(t)$.

Lemma 2.3. *We have for $t \geq 0$ and $\theta \in \mathbb{R}$.*

i) The characteristic function of $Z(t)$ started from $Z(0) = b$,

$$\varphi_t(\theta) = \mathbb{E}\left[e^{i\theta Z(t)} | Z(0) = b\right] = \left(\frac{e^{i\theta(b-1)}e^{-(b-1)t}}{1 - e^{i\theta(b-1)} + e^{i\theta(b-1)}e^{-(b-1)t}}\right)^{\frac{b}{b-1}}. \tag{2.6}$$

ii) We have

$$\mathbb{E}[e^{i\theta Z_m(t)}] = \mathbb{E}\left[\exp\left((1-p)\int_0^t Z_0^{(p)}(t-s)\,(\varphi_s(\theta)-1)\,\mathrm{d}s\right)\right]. \tag{2.7}$$

Proof. Recall that the processes $(Z_i(t - a_i^{(p)}) : t \geq a_i^{(p)})$ for $i \geq 1$ are i.i.d. branching process with the same law as Z with starting value b. Then according to page 109 in Chapter III of Athreya and Ney [17], their characteristic function is given by the expression (2.6). We now observe from the dynamics of $\mathbf{Z}^{(p)}$ that the counting process M has jumps at rate $(1-p)Z_0^{(p)}$. Moreover, conditionally on $Z_0^{(p)}$, the process Z_{m} is a non homogeneous filtered Poisson process whose characteristic function can be written in terms of the characteristic function of Z_i. By extending equation (5.43) of Parzen [74] slightly to allow the underlying Poisson process to be non homogeneous, we obtain

$$\mathbb{E}\left[e^{i\theta Z_{\mathrm{m}}(t)} | \left(Z_0^{(p)}(s) : 0 \leq s \leq t\right)\right] = \exp\left((1-p)\int_0^t Z_0^{(p)}(s)\,(\varphi_{t-s}(\theta)-1)\,\mathrm{d}s\right), \tag{2.8}$$

for $t \geq 0$ and $\theta \in \mathbb{R}$. Hence our claim follows after taking expectations on both sides of the equation and make a simple change of variables. □

We recall some important properties of the branching processes Z and $Z_0^{(p)}$, which will be useful later on. The process

$$W(t) := e^{-(b-1)t}Z(t), \qquad t \geq 0$$

is a nonnegative square-integrable martingale which converges a.s. and in $L^2(\mathbb{P})$, and we write $W(\infty)$ for its terminal value. Furthermore $W(\infty) > 0$ a.s. since Z can not become extinct (we also pointed out that Z never explodes a.s.). Similarly, the process

$$W_0^{(p)}(t) = e^{-(bp-1)t}Z_0^{(p)}(t), \qquad t \geq 0$$

is a martingale which terminal value is denoted by $W_0^{(p)}(\infty)$. In addition, following the proof of Lemma 3 in [32] we have

Lemma 2.4. *It holds that*

$$\lim_{p\to 1, t\to\infty} \mathbb{E}_z\left[\sup_{s\geq t}\left|W_0^{(p)}(s) - W(\infty)\right|^2\right] = 0.$$

In particular, $W_0^{(p)}(\infty)$ converges to $W(\infty)$ in $L^2(\mathbb{P})$ as $p \to 1$.

We next estimate the characteristic function of $Z_{\mathrm{m}}(t)$ given in (2.7), but we still need some additional notation. For $t \geq 0$,

$$I^{(p)}(t) = (1-p)\int_0^t Z_0^{(p)}(t-s)(\varphi_s(u)-1)\mathrm{d}s$$

and

$$I_{\mathrm{m}}^{(p)}(t) = (1-p)W_0^{(p)}(\infty)e^{(b-1)t}\int_0^t e^{-(b-1)s}(\varphi_s(u)-1)\mathrm{d}s,$$

where $u = \theta/\left(\beta c \ln^3 n\right)$ for $\theta \in \mathbb{R}$ and $\beta = b/(b-1)$.

Lemma 2.5. *We have*

$$\lim_{n\to\infty}\left(I^{(p)}(\tau_0(\ln^4 n)) - I_{\mathrm{m}}^{(p)}(\tau_0(\ln^4 n))\right) = 0 \qquad \textit{in probability.}$$

Proof. Define the function

$$I_{\mathrm{a}}^{(p)}(t) = (1-p)W_0^{(p)}(\infty)e^{(bp-1)t}\int_0^t e^{-(bp-1)s}(\varphi_s(u)-1)\mathrm{d}s, \qquad t \geq 0,$$

which is simply obtained by replacing b by bp in the exponential terms of $I_{\mathrm{m}}^{(p)}(t)$. We first prove that

$$\lim_{n\to\infty}\left(I^{(p)}(\tau_0(\ln^4 n)) - I_{\mathrm{a}}^{(p)}(\tau_0(\ln^4 n))\right) = 0 \quad \text{in probability.} \tag{2.9}$$

In this direction, we observe from the triangle inequality that

$$\begin{aligned}&\left|I^{(p)}(\tau_0(\ln^4 n)) - I_{\mathrm{a}}^{(p)}(\tau_0(\ln^4 n))\right| \\ &\quad\leq (1-p)e^{(bp-1)\tau_0(\ln^4 n)}\int_0^{\tau_0(\ln^4 n)} |W_0^{(p)}(\tau_0(\ln^4 n)-s) - W_0^{(p)}(\infty)||\varphi_s(u)-1|e^{-(bp-1)s}\mathrm{d}s.\end{aligned} \tag{2.10}$$

We define

$$A^{(p)} := \frac{3}{2(bp-1)}\sup_{s\geq 0} e^{(bp-1)s/3}|W_0^{(p)}(s) - W_0^{(p)}(\infty)|,$$

and since Lemma 2 in [32] shows that $A^{(p)}$ is bounded in $L^2(\mathbb{P})$, we have by the Markov inequality that

$$\lim_{n\to\infty}\left(\ln^{-\frac{1}{3}} n\right)A^{(p)} = 0 \quad \text{in probability.} \tag{2.11}$$

We set $t_n = (b-1)^{-1}\ln\ln n$ and recall that $\varphi_t(u)$ fulfills (2.6). Hence from the inequality $|e^{ix}-1| \leq 2$ for $x \in \mathbb{R}$, we have that

$$|\varphi_t(u) - 1| \leq 2. \tag{2.12}$$

Then,

$$\begin{aligned}
(1-p)e^{(bp-1)\tau_0(\ln^4 n)} &\int_{\tau_0(\ln^4 n)-t_n}^{\tau_0(\ln^4 n)} |W_0^{(p)}(\tau_0(\ln^4 n)-s) - W_0^{(p)}(\infty)||\varphi_s(u)-1|e^{-(bp-1)s}\mathrm{d}s \\
&\leq 2(1-p)\int_0^{t_n} |W_0^{(p)}(s) - W_0^{(p)}(\infty)|e^{(bp-1)s}\mathrm{d}s \\
&\leq 2(1-p)\left(\ln^{\frac{2}{3}} n\right) A^{(p)}.
\end{aligned} \tag{2.13}$$

On the other hand, from (2.6) and the inequality $|e^{ix}-1| \leq |x|$ for $x \in \mathbb{R}$, we get that

$$|\varphi_t(u) - 1| \leq b|u|e^{(b-1)t}, \tag{2.14}$$

which implies that

$$\begin{aligned}
(1-p)e^{(bp-1)\tau_0(\ln^4 n)} &\int_0^{\tau_0(\ln^4 n)-t_n} |W_0^{(p)}(\tau_0(\ln^4 n)-s) - W_0^{(p)}(\infty)||\varphi_s(u)-1|e^{-(bp-1)s}\mathrm{d}s \\
&\leq (1-p)b|u|e^{(b-1)\tau_0(\ln^4 n)}\int_{t_n}^{\tau_0(\ln^4 n)} |W_0^{(p)}(s) - W_0^{(p)}(\infty)|e^{-(b-1)s}e^{(bp-1)s}\mathrm{d}s \\
&\leq 2b|u|(1-p)\left(\ln^{-\frac{1}{3}} n\right) A^{(p)}e^{(b-1)\tau_0(\ln^4 n)}.
\end{aligned} \tag{2.15}$$

We recall that $u = \theta/\left(\beta c \ln^3 n\right)$, and $p = p_n$ is given by (2.1). Then from (2.10), (2.13), and (2.15) follow that

$$\left| I^{(p)}(\tau_0(\ln^4 n)) - I_{\mathrm{a}}^{(p)}(\tau_0(\ln^4 n)) \right| \leq 2\left(c + (b-1)|\theta|\left(\ln^{-4} n\right)e^{(b-1)\tau_0(\ln^4 n)}\right)\left(\ln^{-\frac{1}{3}} n\right) A^{(p)}.$$

We observe that since $Z_0^{(p)}(\tau_0(\ln^4 n)) = (b-1)\lfloor \ln^4 n \rfloor + 1$, Lemma 2.4 ensures that

$$\lim_{n\to\infty}\left(\ln^{-4} n\right)e^{(b-1)\tau_0(\ln^4 n)} = \frac{b-1}{W(\infty)} \qquad \text{in probability}, \tag{2.16}$$

where $W(\infty)$ is strictly positive almost surely. Thus, we deduce (2.9) from (2.11) by letting $n \to \infty$ in the last inequality.

Next, we prove that

$$\lim_{n\to\infty}\left(I_{\mathrm{m}}^{(p)}(\tau_0(\ln^4 n)) - I_{\mathrm{a}}^{(p)}(\tau_0(\ln^4 n))\right) = 0 \qquad \text{in probability}, \tag{2.17}$$

by proceeding similarly as the proof of (2.9). We observe for the triangle inequality that

$$\begin{aligned}
&\left| I_{\mathrm{m}}^{(p)}(\tau_0(\ln^4 n)) - I_{\mathrm{a}}^{(p)}(\tau_0(\ln^4 n)) \right| \\
&\quad\leq (1-p)W_0^{(p)}(\infty)\int_0^{\tau_0(\ln^4 n)} \left|1 - e^{b(1-p)(s-\tau_0(\ln^4 n))}\right| |\varphi_s(u)-1|e^{-(b-1)(s-\tau_0(\ln^4 n))}\mathrm{d}s.
\end{aligned} \tag{2.18}$$

We set $t_n = (b-1)^{-1} \ln\ln n$ again. Hence from the inequality (2.12) we have that

$$
\begin{aligned}
&(1-p)W_0^{(p)}(\infty) \int_{\tau_0(\ln^4 n)-t_n}^{\tau_0(\ln^4 n)} \left|1 - e^{b(1-p)(s-\tau_0(\ln^4 n))}\right| |\varphi_s(u) - 1| e^{-(b-1)(s-\tau_0(\ln^4 n))} \mathrm{d}s \\
&\quad \leq 2(1-p)W_0^{(p)}(\infty) \int_{\tau_0(\ln^4 n)-t_n}^{\tau_0(\ln^4 n)} \left|1 - e^{b(1-p)(s-\tau_0(\ln^4 n))}\right| e^{-(b-1)(s-\tau_0(\ln^4 n))} \mathrm{d}s \\
&\quad = 2(1-p)W_0^{(p)}(\infty) \int_0^{t_n} \left|1 - e^{-b(1-p)s}\right| e^{(b-1)s} \mathrm{d}s.
\end{aligned}
$$

Then by making the change of variables $x = e^{(b-1)s}$, we get that

$$
\begin{aligned}
&(1-p)W_0^{(p)}(\infty) \int_{\tau_0(\ln^4 n)-t_n}^{\tau_0(\ln^4 n)} \left|1 - e^{b(1-p)(s-\tau_0(\ln^4 n))}\right| |\varphi_s(u) - 1| e^{-(b-1)(s-\tau_0(\ln^4 n))} \mathrm{d}s \\
&\quad \leq \frac{2}{b-1}(1-p)W_0^{(p)}(\infty) \int_1^{\ln n} \left(1 - x^{-\beta(1-p)}\right) \mathrm{d}x \\
&\quad \leq \frac{2}{b-1}(1-p)W_0^{(p)}(\infty) \left(1 - (\ln n)^{-\beta(1-p)}\right) \ln n. \qquad (2.19)
\end{aligned}
$$

On the other hand, from the inequality (2.14) we have that

$$
\begin{aligned}
&(1-p)W_0^{(p)}(\infty) \int_0^{\tau_0(\ln^4 n)-t_n} \left|1 - e^{b(1-p)(s-\tau_0(\ln^4 n))}\right| |\varphi_s(u) - 1| e^{-(b-1)(s-\tau_0(\ln^4 n))} \mathrm{d}s \\
&\quad \leq (1-p)b|u|e^{(b-1)\tau_0(\ln^4 n)} W_0^{(p)}(\infty) \int_{t_n}^{\tau_0(\ln^4 n)} \left(1 - e^{-b(1-p)s}\right) \mathrm{d}s \\
&\quad \leq \frac{b^2}{2}(1-p)^2 |u| e^{(b-1)\tau_0(\ln^4 n)} W_0^{(p)}(\infty)(\tau_0(\ln^4 n))^2. \qquad (2.20)
\end{aligned}
$$

Recall that $u = \theta / \left(\beta c \ln^3 n\right)$, and $p = p_n$ is given by (2.1). Then from (2.18), (2.19), and (2.20) follow that

$$
\begin{aligned}
&\left|I_{\mathrm{m}}^{(p)}(\tau_0(\ln^4 n)) - I_{\mathrm{a}}^{(p)}(\tau_0(\ln^4 n))\right| \\
&\quad \leq 2b(b-1)cW_0^{(p)}(\infty) \left(1 - e^{-\beta c \frac{\ln\ln n}{\ln n}} + |\theta|(\tau_0(\ln^4 n))^2 e^{(b-1)\tau_0(\ln^4 n)} \ln^{-5} n\right).
\end{aligned}
$$

We deduce from (2.16) that

$$
\lim_{n\to\infty} \frac{\tau_0(\ln^4 n)}{4(b-1)^{-1}\ln\ln n} = 1 \qquad \text{in probability,}
$$

and since $\lim_{n\to\infty} W_0^{(p)}(\infty) = W(\infty)$ in $L^2(\mathbb{P})$ from Lemma 2.4, we get (2.17) from (2.16) by letting $n \to \infty$ in the last inequality. Finally, our claim follows by combining (2.9) and (2.17). $\square$

We observe that thanks to (2.6), the integral $I_{\mathrm{m}}^{(p)}$ can be computed explicitly.

Lemma 2.6. *We have for $t \geq 0$,*

$$
\int_0^t e^{-(b-1)s}(\varphi_s(u) - 1)\mathrm{d}s = \frac{1 - e^{iu(b-1)}}{(b-1)e^{iu(b-1)}} \left(\beta \ln(1 - e^{iu(b-1)} + e^{iu(b-1)}e^{-(b-1)t}) + \kappa_{b,u}(t)\right),
$$

where

$$\kappa_{b,u}(t) = \sum_{k=2}^{\infty} \frac{(\beta)_k}{k!} \frac{(e^{iu(b-1)} - 1)^{k-1}}{k-1} \left(1 - \frac{1}{(1 - e^{iu(b-1)} + e^{iu(b-1)} e^{-(b-1)t})^{k-1}}\right), \tag{2.21}$$

and $(\cdot)_k$ *is the Pochhammer function.*

Proof. Define the function

$$f(\lambda) = \int_0^t e^{-(b-1)r} \left(\left(\frac{e^{-\lambda(b-1)} e^{-(b-1)r}}{1 - e^{-\lambda(b-1)} + e^{-\lambda(b-1)} e^{-(b-1)r}} \right)^{\beta} - 1 \right) \mathrm{d}r, \quad \lambda \geq 0.$$

Hence by setting $x = 1 - e^{-\lambda(b-1)} + e^{-\lambda(b-1)} e^{-(b-1)r}$ and $y_\lambda = e^{-\lambda(b-1)}$ for convenience, we have that

$$f(\lambda) = \frac{1}{(b-1)y_\lambda} \int_{1-y_\lambda+y_\lambda e^{-(b-1)t}}^{1} \left(\left(\frac{x + y_\lambda - 1}{x} \right)^{\beta} - 1 \right) \mathrm{d}x.$$

Moreover, using a well-known extension of Newton's binomial formula, we get

$$f(\lambda) = \frac{1}{(b-1)y_\lambda} \sum_{k=1}^{\infty} \frac{(\beta)_k}{k!} (y_\lambda - 1)^k \int_{1-y_\lambda+y_\lambda e^{-(b-1)t}}^{1} x^{-k} \mathrm{d}x,$$

where $(\cdot)_k$ is the Pochhammer function. Note that the series converges absolutely since $\beta > 0$ and $|y_\lambda - 1|/x \leq 1$, for $1 - y_\lambda + y_\lambda e^{-(b-1)t} \leq x \leq 1$. Then straightforward calculations yield

$$f(\lambda) = \frac{1 - y_\lambda}{(b-1)y_\lambda} \left(\beta \ln(1 - y_\lambda + y_\lambda e^{-(b-1)t}) + \kappa'_{b,\lambda}(t) \right),$$

where

$$\kappa'_{b,\lambda}(t) = \sum_{k=2}^{\infty} \frac{(\beta)_k}{k!} \frac{(y_\lambda - 1)^{k-1}}{k-1} \left(1 - \frac{1}{(1 - y_\lambda + y_\lambda e^{-(b-1)t})^{k-1}} \right).$$

We note that the function f allows an analytic extension to $\{\lambda \in \mathbb{C} : \mathrm{Re}\, \lambda \geq 0\}$. Then by taking into account the principal branch of the complex logarithm, we conclude that

$$f(-iu) = \int_0^t e^{-(b-1)r} \left(\left(\frac{e^{iu(b-1)} e^{-(b-1)r}}{1 - e^{iu(b-1)} + e^{iu(b-1)} e^{-(b-1)r}} \right)^{\beta} - 1 \right) \mathrm{d}r,$$

and our assertion follows by observing that $\kappa'_{b,\lambda}(t) = \kappa_{b,u}(t)$ when $\lambda = -iu$. $\square$

We are now able to establish Proposition 2.1.

Proof of Proposition 2.1. Fix $\theta \in \mathbb{R}$ and define $m_n = \beta c \ln^3 n$. From the identity (2.5) and Lemma 2.3, the characteristic function of $m_n^{-1}\Delta_{0,n} - (\beta - 1)^{-1} \ln m_n$ is given by

$$\mathbb{E}\left[e^{i\theta(m_n^{-1}\Delta_{0,n} - (\beta-1)^{-1}\ln m_n)}\right] = \mathbb{E}\left[\exp\left(I^{(p)}(\tau_0(\ln^4 n)) - i\theta(\beta-1)^{-1}\ln m_n\right)\right].$$

Recall that by Lemma 2.5 we have

$$\lim_{n\to\infty}\left(I^{(p)}(\tau_0(\ln^4 n)) - I^{(p)}_{\mathrm{m}}(\tau_0(\ln^4 n))\right) = 0 \quad \text{in probability.}$$

Then, we must verify that

$$\begin{aligned}&\lim_{n\to\infty}\left(I^{(p)}_{\mathrm{m}}(\tau_0(\ln^4 n)) - i\theta(\beta-1)^{-1}\ln m_n\right)\\ &\quad= -i\theta(\beta-1)^{-1}\left(\kappa_\beta - 1 + \frac{1}{\beta}\right) - i\theta(\beta-1)^{-1}\ln|(\beta-1)^{-1}\theta| - \frac{1}{2}\pi|(\beta-1)^{-1}\theta| \end{aligned} \tag{2.22}$$

in probability. In this direction, we define $y_u = e^{iu(b-1)}$ for convenience and recall also from Lemma 2.6 that

$$\begin{aligned}&I^{(p)}_{\mathrm{m}}(\tau_0(\ln^4 n))\\ &= (1-p)W_0^{(p)}(\infty)e^{(b-1)\tau_0(\ln^4 n)}\frac{1-y_u}{(b-1)y_u}\left(\beta\ln(1-y_u+y_u e^{-(b-1)\tau_0(\ln^4 n)}) + \kappa_{b,u}(\tau_0(\ln^4 n))\right),\end{aligned}$$

where $\kappa_{b,u}(\tau_0(\ln^4 n))$ is defined in (2.21) and $u = \theta/\left(\beta c \ln^3 n\right)$. We know from Lemma 2.4 that

$$\lim_{n\to\infty} e^{-(b-1)\tau_0(\ln^4 n)}Z(\tau_0(\ln^4 n)) = \lim_{n\to\infty} e^{-(bp-1)\tau_0(\ln^4 n)}Z_0^{(p)}(\tau_0(\ln^4 n)) = W(\infty)$$

in probability, and $\lim_{n\to\infty} W_0^{(p)}(\infty) = W(\infty)$ in $L^2(\mathbb{P})$. Hence since $p = p_n$ fulfilled (2.1), we deduce that

$$\lim_{n\to\infty}\frac{(1-p)W_0^{(p)}(\infty)e^{(b-1)\tau_0(\ln^4 n)}}{(b-1)c\ln^3 n} = 1 \quad \text{in probability.}$$

On the other hand,

$$y_u = 1 + O\left(\frac{1}{m_n}\right) \quad \text{and} \quad m_n(1-y_u) = -i\theta(b-1) + O\left(\frac{1}{m_n}\right),$$

and since $b-1 = (\beta-1)^{-1}$, we deduce that

$$\begin{aligned}&\lim_{n\to\infty}\left(I^{(p)}_{m}(\tau_0(\ln^4 n)) - i\theta(\beta-1)^{-1}\ln m_n\right)\\ &\qquad= -i\theta(\beta-1)^{-1}\ln(-i\theta(\beta-1)^{-1}) - i\theta(\beta-1)^{-1}\left(\kappa_\beta - 1 + \frac{1}{\beta}\right)\end{aligned}$$

in probability, which implies (2.22). Finally, we observe from (2.8) and the modulus inequality for conditional expectation that

$$\left|\exp\left(I^{(p)}(\tau_0(\ln^4 n)) - i\theta(\beta-1)^{-1}\ln m_n\right)\right| \le 1.$$

Therefore, by the dominated convergence theorem we conclude that the Fourier transform of $m_n^{-1}\Delta_n^0 - (\beta-1)^{-1}\ln m_n$ converges pointwise as n tends to infinity to the continuous function

$$\theta \mapsto \exp\left(-i\theta(\beta-1)^{-1}\left(\kappa_\beta - 1 + \frac{1}{\beta}\right) - i\theta(\beta-1)^{-1}\ln|(\beta-1)^{-1}\theta| - \frac{1}{2}\pi|(\beta-1)^{-1}\theta|\right),$$

and then our claim follows for the continuity theorem for Fourier transforms. □

We now turn our attention to the proof of Lemma 2.2.

We imagine that we begin our observation of the system of branching processes with rare mutations $\mathbf{Z}^{(p)}$ once it has reached the size $(b-1)\lfloor \ln^4 n \rfloor + 1$, that is, from the time $\tau(\ln^4 n)$. We thus write $Z' = (Z'(t) : t \geq 0)$ for the process of the total size of the population started from $Z'(0) = (b-1)\lfloor \ln^4 n \rfloor + 1$, which has the same law of the Yule process Z described at the beginning of Section 2.2. We introduce the time

$$\tau'(\ln^4 n) = \inf\{t \geq 0 : Z'(t) = Z(\tau_0(\ln^4 n))\},$$

at which it hits $Z(\tau_0(\ln^4 n))$. Equivalently, $\tau'(\ln^4 n)$ is the time needed to have a population with the ancestral type of size $(b-1)\lfloor \ln^4 n \rfloor + 1$. We shall first estimate this quantity.

Lemma 2.7. *We have*

$$\lim_{n\to\infty} \tau'(\ln^4 n) = 0 \qquad \textit{in probability.}$$

Proof. We know that

$$\lim_{n\to\infty} e^{-(b-1)\tau(\ln^4 n)} Z(\tau(\ln^4 n)) = W(\infty) \qquad \text{in probability.}$$

By definition of the time $\tau(\ln^4 n)$, we have $Z(\tau(\ln^4 n)) = (b-1)\lfloor \ln^4 n \rfloor + 1$, hence

$$\lim_{n\to\infty} \frac{\tau(\ln^4 n)}{4(b-1)^{-1} \ln \ln n} = 1 \qquad \text{in probability.}$$

On the other hand, from Lemma 2.4 we have that

$$\lim_{n\to\infty} e^{-(bp-1)\tau_0(\ln^4 n)} Z_0^{(p)}(\tau_0(\ln^4 n)) = W(\infty) \qquad \text{in probability.}$$

Recall that $p = p_n$ is given by (2.1), and observe that $Z_0^{(p)}(\tau_0(\ln^4 n)) = (b-1)\lfloor \ln^4 n \rfloor + 1$. Hence

$$\lim_{n\to\infty} e^{-(bp-1)(\tau(\ln^4 n) - \tau_0(\ln^4 n))} = 1 \qquad \text{in probability,}$$

and our claim follows from the identity $\tau'(\ln^4 n) = \tau_0(\ln^4 n) - \tau(\ln^4 n)$. □

We observe that the population at time $\tau(\ln^4 n)$ when we start our observation consists of Δ_n mutants and $(b-1)\lfloor \ln^4 n \rfloor + 1 - \Delta_n$ individuals of the ancestral type. Then, we write $Z_0'^{(p)} = (Z_0'^{(p)}(t) : t \geq 0)$ for the process that counts the number of individuals with the ancestral type, which has the same law as $Z_0^{(p)}$ but starting from $Z_0'^{(p)}(0) = (b-1)\lfloor \ln^4 n \rfloor + 1 - \Delta_n$. We recall that

$$W'(t) := e^{-(b-1)t} Z'(t) \qquad \text{and} \qquad W_0'^{(p)}(t) := e^{-(bp-1)t} Z'^{(p)}(t), \qquad t \geq 0$$

are nonnegative square-integrable martingales which converge a.s. and in $L^2(\mathbb{P})$.

Proof of Lemma 2.2. An application of Doob's inequality (see, e.g., Equation (6) in [75]) shows for all $\eta > 0$ that

$$\lim_{n\to\infty} \mathbb{P}\left(\left|e^{-(b-1)\tau'(\ln^4 n)} Z'(\tau'(\ln^4 n)) - Z'(0)\right| > \eta \ln^3 n\right) = 0$$

and using the fact that $Z_0'^{(p)}(0) \leq (b-1)\lfloor \ln^4 n \rfloor + 1$, we also get

$$\lim_{n\to\infty} \mathbb{P}\left(\left|e^{-(bp-1)\tau'(\ln^4 n)} Z_0'^{(p)}(\tau'(\ln^4 n)) - Z_0'^{(p)}(0)\right| > \eta \ln^3 n\right) = 0.$$

Then, since $\Delta_{0,n} = Z'(\tau'(\ln^4 n)) - (b-1)\lfloor \ln^4 n \rfloor - 1$, $\Delta_n = (b-1)\lfloor \ln^4 n \rfloor + 1 - Z_0'^{(p)}(0)$, and $Z'(0) = (b-1)\lfloor \ln^4 n \rfloor + 1$, one readily gets

$$\begin{aligned}
&\Delta_n - \Delta_{0,n} \\
&\quad = Z'(\tau'(\ln^4 n))\left(e^{-(b-1)\tau'(\ln^4 n)} - 1\right) - Z_0'^{(p)}(\tau'(\ln^4 n))\left(e^{-(bp-1)\tau'(\ln^4 n)} - 1\right) + o(\ln^3 n)
\end{aligned}$$

in probability. We next note from Lemma 2.7 that

$$Z_0'^{(p)}(\tau'(\ln^4 n))\left(e^{(1-p)\tau'(\ln^4 n)} - 1\right) = o(\ln^3 n) \quad \text{in probability},$$

which yields

$$\Delta_n - \Delta_{0,n} = \left(W'(\tau'(\ln^4 n)) - W_0'^{(p)}(\tau'(\ln^4 n))\right)\left(1 - e^{(b-1)\tau'(\ln^4 n)}\right) + o(\ln^3 n)$$

in probability. Since by Lemma 2.2 we have that

$$\lim_{n\to\infty}\left(1 - e^{(b-1)\tau'(\ln^4 n)}\right) = 0 \quad \text{in probability},$$

we must verify that

$$W'(\tau'(\ln^4 n)) - W_0'^{(p)}(\tau'(\ln^4 n)) = o(\ln^3 n) \quad \text{in probability},$$

in order to get the result of Lemma 2.2. We observe from properties of square-integrable martingales that

$$\mathbb{E}\left[\left(W'(\tau'(\ln^4 n))\right)^2\right] = \mathbb{E}\left[\left[W'\right]_{\tau'(\ln^4 n)}\right]$$

where

$$\left[W'\right]_t = \sum_{0\leq s\leq t} e^{-2(b-1)s}|Z'(s) - Z'(s-)|^2 \quad \text{for} \quad t \geq 0.$$

A straightforward calculation shows that the compensator of jump process $[W']$ is

$$\langle W' \rangle_t = (b-1)^2 \int_0^t e^{-2(b-1)s} Z'(s)\mathrm{d}s \quad \text{for} \quad t \geq 0,$$

that is $[W']_t - \langle W' \rangle_t$ is a local martingale. Thus,

$$\mathbb{E}\left[\left(W'(\tau'(\ln^4 n))\right)^2\right] = \mathbb{E}\left[\langle W' \rangle_{\tau'(\ln^4 n)}\right] \leq \mathbb{E}\left[\langle W' \rangle_\infty\right] = (b-1)((b-1)\lfloor \ln^4 n \rfloor + 1).$$

Hence by the Markov inequality we have that

$$W'(\tau'(\ln^4 n)) = o(\ln^3 n) \qquad \text{in probability.}$$

Similarly one gets

$$W_0'^{(p)}(\tau'(\ln^4 n)) = o(\ln^3 n) \qquad \text{in probability,}$$

from where our claim follows. □

2.2.2 The spread of fluctuations

The purpose here is to resume the growth of the system of branching processes with rare mutation from the size $(b-1)\lfloor \ln^4 n \rfloor + 1$ to the size $(b-1)n+1$ and observe that the germ of the fluctuations Δ_n spreads regularly. In this direction, we proceed similarly as the last part of the preceding section. We recall that Z' denotes the process of the total population started from $Z'(0) = (b-1)\lfloor \ln^4 n \rfloor + 1$. We consider

$$\tau'(n) = \inf\{t \geq 0 : Z'(t) = (b-1)n + 1\},$$

the time needed for the total population to reach size $(b-1)n+1$. Hence, in the notation of Theorem 2.2

$$G_n = Z_0'^{(p)}(\tau'(n)),$$

where as the previous section, we write $Z_0'^{(p)}$ for the process that counts the number of individuals with the ancestral type starting from $Z_0'^{(p)}(0) = (b-1)\lfloor \ln^4 n \rfloor + 1 - \Delta_n$.

We have now all the ingredients to establish Theorem 2.2.

Proof of Theorem 2.2. Again from the estimate of Equation (6) in [75], we get for all $\eta > 0$ that

$$\lim_{n\to\infty} \mathbb{P}\left(\left|((b-1)n+1)e^{-(b-1)\tau'(n)} - (b-1)\ln^4 n - 1\right| > \eta \ln^3 n\right) = 0,$$

this yields

$$e^{(b-1)\tau'(n)} = \frac{n}{\ln^4 n} + o\left(\frac{1}{\ln n}\right) \qquad \text{in probability.} \tag{2.23}$$

On the other hand, using the fact that $Z_0'^{(p)}(0) \leq (b-1)\lfloor \ln^4 n \rfloor + 1$, we also get

$$\lim_{n\to\infty} \mathbb{P}\left(\left|e^{-(bp-1)\tau'(n)} Z_0'^{(p)}(\tau'(n)) - Z_0'^{(p)}(0)\right| > \eta \ln^3 n\right) = 0,$$

and deduce that

$$G_n = e^{(bp-1)\tau'(n)}((b-1)\ln^4 n - \Delta_n) + o\left(\frac{n}{\ln n}\right) \qquad \text{in probability.}$$

Next, it is convenient to apply Skorokhod's representation theorem and assume that the weak convergence in Corollary 2.1 holds almost surely. Hence

$$G_n = e^{(bp-1)\tau'(n)}\left((b-1)\ln^4 n - \frac{\beta}{\beta-1}c\ln^3 n\left(3\ln\ln n + \left(\mathcal{Z}_{c,\beta} + 1 - \frac{1}{\beta}\right)\right)\right) + o\left(\frac{n}{\ln n}\right)$$

in probability. We next note from (2.23) that

$$e^{(bp-1)\tau'(n)} = e^{-\beta c}\frac{n}{\ln^4 n} + 4\beta c e^{-\beta c} n \frac{\ln\ln n}{\ln^5 n} + o\left(\frac{n}{\ln^5 n}\right)$$

in probability. Recall that $\beta = b/(b-1)$, then

$$G_n = \frac{1}{\beta-1}e^{-\beta c}n + \frac{\beta}{\beta-1}ce^{-\beta c}n\frac{\ln\ln n}{\ln n} - \frac{\beta}{\beta-1}ce^{-\beta c}\frac{n}{\ln n}\left(\mathcal{Z}_{c,\beta} + 1 - \frac{1}{\beta}\right) + o\left(\frac{n}{\ln n}\right)$$

in probability, which completes the proof. □

2.3 Proof of Theorem 2.1

Our approach is based in the introduction of a continuous version of the construction of a b-ary recursive tree that enables us to superpose Bernoulli bond percolation dynamically in the tree structure. We begin at time 0 from the tree with just one internal vertex which corresponds to the root having b external vertices. Once the random tree with size $n \geq 1$ has been constructed, we equip each of the $(b-1)n+1$ external vertices with independent exponential random variables ζ_i of parameter 1. Then, after a waiting time equal to $\min_{i\in\{1,\dots,(b-1)n+1\}} \zeta_i$, one of the external vertices is chosen uniformly at random and is replaced it by the internal vertex $n+1$ to which b new leaves are attached. We observe that $\min_{i\in\{1,\dots,(b-1)n+1\}} \zeta_i$ is exponentially distributed with parameter $(b-1)n+1$.

We denote by $T(t)$ the tree which has been constructed at time $t \geq 0$, and by $|T(t)|$ its size, i.e. the number of internal vertices. The process of the size $(|T(t)| : t \geq 0)$ is clearly Markovian and if we define

$$\gamma(n) = \inf\{t \geq 0 : |T(t)| = n\}, \qquad n \geq 1,$$

then $T(\gamma(n))$ is a version of the b-ary recursive tree of size n, T_n. However for our purpose it will be more convenient work with the process Y defined by

$$Y(t) = (b-1)|T(t)| + 1, \quad t \geq 0 \tag{2.24}$$

with starting value $Y(0) = b$. It should be clear that Y is a Yule process as described in Section 2.2, i.e. it has jumps of size $b-1$ and unit birth rate per unit population size. We also point out that the process Y gives us the number of external vertices on the tree.

We next superpose Bernoulli bond percolation with parameter $p = p_n$ defined in (2.1) to the growth algorithm in continuous time of the b-ary recursive tree. We follow the approach developed by Bertoin and Uribe Bravo [32] but with a slight modification. We draw an independent Bernoulli random variable ϵ_p with parameter p, each time an internal vertex is inserted. The edge which connects this new internal vertex is cut at its midpoint when $\epsilon_p = 0$ and remains intact otherwise. This disconnects the tree into connected clusters which motivates the following. We write $T^{(p)}(t)$ for the resulting combinatorial structure at time t. So, the percolation clusters of $T(t)$ are the connected components by a path of intact edges of $T^{(p)}(t)$.

Let $T_0^{(p)}(t)$ be the subtree that contains the root. We write $H_0^{(p)}(t)$ for the number of half-edges pertaining to the root cluster at time t. So that, its number of external vertices is given by

$$Y_0^{(p)}(t) = (b-1)|T_0^{(p)}(t)| + 1 - H_0^{(p)}(t).$$

We are now be able to observe the connection with the system of branching processes with rare mutations described in the preceding section. It should be plain from the construction that the size of the root-cluster at time t, i.e. $Y_0^{(p)}(t)$, of $T(t)$ after percolation with parameter p, coincides with the number of individuals with the ancestral type $Z_0^{(p)}(t)$ in the system $\mathbf{Z}^{(p)}$ of branching processes with rare mutations of Section 2.2. In fact, we already mentioned that the process Y has the same random evolution as the process of the total size in the system Z. Recall that the algorithm for constructing a b-ary recursive tree is run until the time

$$\gamma(n) = \inf\{t \geq 0 : |T(t)| = n\} = \inf\{t \geq 0 : Y(t) = (b-1)n + 1\}$$

when the structure has n internal vertices. Then, the size $C_0^{(p)}$ of the percolation cluster containing the root when the tree has n internal vertices satisfies

$$C_0^{(p)} = |T_0^{(p)}(\gamma(n))|.$$

In addition, it should be plain that

$$Y_0^{(p)}(\gamma(n)) = (b-1)C_0^{(p)} + 1 - H_0^{(p)}(\gamma(n)),$$

coincides with the number of individuals with the ancestral type in the branching system $\mathbf{Z}^{(p)}$, at time when the total population reaches the size $(b-1)n+1$, i.e. G_n, according to the notation of Theorem 2.2. Hence in order to establish Theorem 2.1, it is sufficient to get an estimate of the number of half-edges pertaining to the root-subtree at time $\gamma(n)$.

Lemma 2.8. *We have*

$$\lim_{n\to\infty} \frac{\ln n}{n} H_0^{(p)}(\gamma(n)) = ce^{-\beta c} \quad \textit{in probability.}$$

Proof. We observe that the processes

$$H_0^{(p)}(t) - (1-p)\int_0^t Y_0^{(p)}(s)\mathrm{d}s \quad \text{and} \quad Y_0^{(p)}(t) - (bp-1)\int_0^t Y_0^{(p)}(s)\mathrm{d}s, \quad t \geq 0$$

are martingales. Thus,

$$L^{(p)}(t) := H_0^{(p)}(t) - \frac{1-p}{bp-1} Y_0^{(p)}(t), \qquad t \geq 0$$

is also a martingale. Observe that since $p = p_n$ satisfies (2.1), for n large enough such that $2/(b+1) \leq p \leq 1$, its jumps $|L^{(p)}(t) - L^{(p)}(t-)|$ have size at most b. Since there are at most n jumps up to time $\gamma(n)$, the bracket of $L^{(p)}$ can be bounded by $[L^{(p)}]_{\gamma(n)} \leq b^2 n$. Hence we have

$$\lim_{n\to\infty} \mathbb{E}\left(\left| \frac{\ln n}{n} L^{(p)}(\gamma(n)) \right|^2 \right) = 0. \tag{2.25}$$

On the other hand, we know from Lemma 2.4 that

$$\lim_{n\to\infty} e^{-(b-1)\gamma(n)} Y(\gamma(n)) = \lim_{n\to\infty} e^{-(bp-1)\gamma(n)} Y_0^{(p)}(\gamma(n)) = W(\infty) \quad \text{in probability}$$

which implies that

$$\lim_{n\to\infty} \frac{Y_0^{(p)}(\gamma(n))}{n} = (b-1)e^{-\beta c} \quad \text{in probability,}$$

and the result follows readily from (2.25), the above limit and the fact that $1-p = o(1)$. $\square$

Therefore, from the identity

$$C_0^{(p)} = \frac{Y_0^{(p)}(\gamma(n)) - 1 + H_0^{(p)}(\gamma(n))}{(b-1)},$$

Theorem 2.2 applies to $Y_0^{(p)}(\gamma(n))$ and Lemma 2.8 yields the result of Theorem 2.1.

2.4 Percolation on scale-free trees

We conclude this work by showing that the approach used in the proof of Theorem 2.1 can be also applied to study percolation on scale-free random trees. Recall that they form a family of random trees on a set of ordered vertices, say $\{0, 1, \dots, n\}$, that grow following a preferential attachment algorithm. Specifically, fix a parameter $a \in (-1, \infty)$, and start for $n = 1$ from the tree $T_1^{(a)}$ on $\{0, 1\}$ which has a single edge connecting 0 and 1. Suppose that $T_n^{(a)}$ has been constructed for some $n \geq 2$, and for every $i \in \{0, 1, \dots, n\}$, denote by $d_n(i)$ the degree of the vertex i in $T_n^{(a)}$. Then conditionally given $T_n^{(a)}$, the tree $T_{n+1}^{(a)}$ is built by adding an edge between the new vertex $n+1$ and a vertex v_n in $T_n^{(a)}$ chosen at random according to the law

$$\mathbb{P}\left(v_n = i | T_n^{(a)}\right) = \frac{d_n(i) + a}{2n + a(n+1)}, \qquad i \in \{0, 1, \dots, n\}.$$

Clearly, the preceding expression defines a probability measure since the sum of the degrees of a tree with $n+1$ vertices equals $2n$. Note also that when one lets $a \to \infty$ the algorithm yields an uniform recursive tree since v_n becomes uniformly distributed on $\{0, 1, \dots, n\}$. We then perform Bernoulli bond percolation with parameter given by (2.1), i.e. $p_n = 1 - c/\ln n$, where $c > 0$ is fixed. It has been observed by Bertoin and Uribe Bravo [32] that this choice of the percolation parameter corresponds to the supercritical regime. More precisely, the size of the cluster $\Gamma_n^{(\alpha)}$ containing the root satisfies

$$\lim_{n\to\infty} n^{-1}\Gamma_n^{(\alpha)} = e^{-\alpha c} \qquad \text{in probability,}$$

where $\alpha = (1+a)/(2+a)$. We are interested in the fluctuations of $\Gamma_n^{(\alpha)}$, and show that an analogous result to Theorem 2.1 holds for large scale-free random trees.

Theorem 2.3. *Set $\alpha = (1+a)/(2+a)$, and assume that the percolation parameter p_n is given by (2.1). Then as $n \to \infty$, there is the weak convergence*

$$\left(n^{-1}\Gamma_n^{(\alpha)} - e^{-\alpha c}\right)\ln n - \alpha c e^{-\alpha c}\ln\ln n \Rightarrow -\alpha c e^{-\alpha c}\mathcal{Z}'_{c,\alpha}$$

where

$$\mathcal{Z}'_{c,\alpha} = \mathcal{Z} - \kappa'_\alpha + \ln(\alpha c) \tag{2.26}$$

with $\mathcal{Z}$ the continuous Luria-Delbrück distribution and

$$\kappa'_\alpha = 1 - \frac{1}{\alpha} + \frac{1}{\alpha}\sum_{k=2}^{\infty}\frac{(\alpha)_k}{k!}\frac{(-1)^k}{k-1}.$$

We now focus on the proof of Theorem 2.3. We follow the route used in the proof of Theorem 2.1, and we analyze a system of branching processes with rare neutral mutations. We point out that in order to avoid repetitions, some technical details will be skipped.

We start by considering a pure birth branching process $Z^{(a)} = (Z^{(a)}(t) : t \geq 0)$ in continuous space, that has only jumps of size $2+a$, and with unit birth rate per unit population size. We shall be mainly interested in a class of population systems which arises by incorporating mutations to the preceding branching process. More precisely, we describe the evolution of such a system by a process $\mathbf{Z}^{(p,a)} = (\mathbf{Z}^{(p,a)} : t \geq 0)$, where for each $t \geq 0$, $\mathbf{Z}^{(p,a)}(t) = (Z_0^{(p,a)}(t), Z_1^{(a)}(t), \dots)$ is a collection of nonnegative variables. At the initial time, all the sub-populations $Z_i^{(a)}(0)$ of type $i \geq 1$ are taken to be equal to zero, and $Z_0^{(p,a)}(0) = 2+2a$ which is the size of the ancestral (or clone) population. We consider that at rate p per unit population size, the clone population produces $2+a$ new clones, and that at rate $1-p$ per unit population size, they always create a single mutant population of a new type of size $1+a$. The new mutant populations behave as the process $Z^{(a)}$ but starting from $1+a$. Clearly, the sum over all sub-populations

$$Z^{(a)}(t) = Z_0^{(p,a)}(t) + \sum_{i\geq 1} Z_i^{(a)}(t), \qquad t \geq 0,$$

evolves as the pure birth branching process described at the beginning of this paragraph.

We next observe the growth of the system of branching process $\mathbf{Z}^{(p,a)}$ until the time

$$\tau^{(a)}(\ln^4 n) = \inf\{t \geq 0 : Z^{(a)}(t) = (2+a)\lfloor \ln^4 n \rfloor + a\},$$

which is when the total size of the population reaches $(2+a)\lfloor \ln^4 n \rfloor + a$. Our first purpose is to estimate precisely the number $\Delta_n^{(\alpha)}$ of mutants at this time. This stage corresponds to the analysis of the germ, and approach line will be similar to that in Section 2.2.1. In this direction, it will be useful to study the number of mutants $\Delta_{0,n}^{(\alpha)}$ at time

$$\tau_0^{(a)}(\ln^4 n) = \inf\{t \geq 0 : Z_0^{(p,a)}(t) = (2+a)\lfloor \ln^4 n \rfloor + a\}$$

whose distribution is easier to estimate than that of $\Delta_n^{(\alpha)}$. We shall establish the following limit theorem in law that is equivalent to the Proposition 2.1.

Proposition 2.2. *As $n \to \infty$, there is the weak convergence*

$$\frac{\Delta_{0,n}^{(\alpha)}}{\ln^3 n} - 3\frac{\alpha}{1-\alpha} c \ln\ln n \Rightarrow \frac{\alpha}{1-\alpha} c \left(\mathcal{Z}'_{c,\alpha} + 1 - \frac{1}{\alpha} \right)$$

where $\mathcal{Z}'_{c,\alpha}$ is the random variable defined in (2.26).

Before proving the Proposition 2.2, it is convenient to introduce the following representation of the total mutant population as we have done in Section 2.2.1. For $i \geq 1$, we write

$$b_i^{(p)} = \inf\{t \geq 0 : Z_i^{(a)}(t) > 0\},$$

for the birth time of the sub-population with type i. Then the processes $(Z_i^{(a)}(t - b_i^{(p)}) : t \geq b_i^{(p)})$ form a sequence of i.i.d. branching processes with the same law as $Z^{(a)}$ but starting value $1+a$, which is independent of the birth-times $(b_i^{(p)})_{i\geq 1}$ and the process $Z_0^{(p,a)}$. Moreover, this sequence is also independent of the process that counts the number of mutation events which is defined by $M^{(a)}(t) = \max\{i \geq 1 : Z_i^{(a)}(t) > 0\}$. Thus, since the jump times of $M^{(a)}$ are in fact $b_1^{(p)} < b_2^{(p)} < \dots$, we can express the total mutant population at time $t \geq 0$ as,

$$Z_{\mathrm{m}}^{(a)}(t) = \sum_{i=1}^{M^{(a)}(t)} Z_i^{(a)}(t - b_i^{(p)}).$$

We observe that for $i \geq 1$, the process $((2+a)^{-1} Z_i^{(a)}(t - b_i^{(p)}) : t \geq b_i^{(p)})$ is a Yule branching process in continuous space with birth rate $2+a$ per unit population size. Then similarly as we obtained the result in Lemma 2.3, we get for $t \geq 0$ and $\theta \in \mathbb{R}$,

$$\mathbb{E}[e^{i\theta Z_{\mathrm{m}}^{(a)}(t)}] = \mathbb{E}\left[\exp\left((1-p)\int_0^t Z_0^{(p,a)}(t-s)(\varphi_s^{(a)}(\theta) - 1)\mathrm{d}s\right)\right] \tag{2.27}$$

where

$$\varphi_t^{(a)}(\theta) = \mathbb{E}\left[e^{i\theta Z^{(a)}(t)} \middle| Z^{(a)}(0) = 1+a\right] = \left(\frac{e^{i\theta(2+a)}e^{-(2+a)t}}{1 - e^{i\theta(2+a)} + e^{i\theta(2+a)}e^{-(2+a)t}}\right)^{\alpha}$$

with $\alpha = (1+a)/(2+a)$.

At this point, the difference between the constants in Theorem 2.1 and Theorem 2.3 must be evident, mostly due to the different behavior of the branching processes associated to the b-ary recursive trees and scale-free random trees. Essentially, the constant κ_β of Theorem 2.1 depends of the characteristic function (2.6) through the computations made in Lemma 2.6, which is clearly distinct from (2.27).

We are now able to establish Proposition 2.2.

Proof of Proposition 2.2. We fix $\theta \in \mathbb{R}$ and define $m_n = \alpha c \ln^3 n$. Since we have the identity $\Delta_{0,n}^{(\alpha)} = Z_{\mathrm{m}}^{(a)}(\tau_0^{(a)}(\ln^4 n))$, it follows from (2.27) that the characteristic function of $m_n^{-1}\Delta_{0,n}^{(\alpha)} - (1-\alpha)^{-1}\ln m_n$ is given by

$$\mathbb{E}\left[e^{i\theta(m_n^{-1}\Delta_{0,n}^{(\alpha)}-(1-\alpha)^{-1}\ln m_n)}\right] = \mathbb{E}\left[\exp\left(I^{(p,a)}(\tau_0^{(a)}(\ln^4 n)) - (1-\alpha)^{-1}\ln m_n\right)\right], \tag{2.28}$$

where

$$I^{(p,a)}(t) = (1-p)\int_0^t Z_0^{(p,a)}(t-s)(\varphi_s^{(a)}(u)-1)\mathrm{d}s \quad \text{for } t \geq 0,$$

and $u = \theta/(\alpha c \ln^3 n)$. Next, a similar computation as the proof of Lemma 2.5 shows that

$$\lim_{n\to\infty}\left(I^{(p,a)}(\tau_0^{(a)}(\ln^4 n)) - I_{\mathrm{m}}^{(p,a)}(\tau_0^{(a)}(\ln^4 n))\right) = 0 \quad \text{in probability,} \tag{2.29}$$

where

$$I_{\mathrm{m}}^{(p,a)}(t) = (1-p)W_0^{(p,a)}(\infty)e^{(2+a)t}\int_0^t e^{-(2+a)s}(\varphi_s^{(a)}(u)-1)\mathrm{d}s \quad \text{for } t \geq 0.$$

and $W_0^{(p,a)}(\infty)$ is the terminal value of the martingale $W_0^{(p,a)}(t) = e^{-(1+p(1+a))t}Z_0^{(p,a)}(t)$. Moreover, the integral of the previous expression can be computed explicitly,

$$\int_0^t e^{-(2+a)s}(\varphi_s^{(a)}(u)-1)\mathrm{d}s = \frac{1-e^{iu(2+a)}}{(2+a)e^{iu(2+a)}}\left(\alpha\ln(1-e^{iu(2+a)}+e^{iu(2+a)}e^{-(2+a)t}) + \kappa'_{\alpha,u}(t)\right),$$

with

$$\kappa'_{\alpha,u}(t) = \sum_{k=2}^{\infty}\frac{(\alpha)_k}{k!}\frac{(e^{iu(2+a)}-1)^{k-1}}{k-1}\left(1-\frac{1}{(1-e^{iu(2+a)}+e^{iu(2+a)}e^{-(2+a)t})^{k-1}}\right).$$

Hence from Lemma 3 in [32] (which is the analog of Lemma 2.4), we conclude after some computations that

$$\begin{aligned}\lim_{n\to\infty}&\left(I_{\mathrm{m}}^{(p,a)}(\tau_0^{(a)}(\ln^4 n)) - i\theta(1-\alpha)^{-1}\ln m_n\right)\\ &= -i\theta(1-\alpha)^{-1}\left(\kappa'_\alpha - 1 + \frac{1}{\alpha}\right) - i\theta(1-\alpha)^{-1}\ln|(1-\alpha)^{-1}\theta| - \frac{1}{2}\pi|(1-\alpha)^{-1}\theta|\end{aligned}$$

in probability, and our claim follows from (2.29), by letting $n\to\infty$ in (2.28). □

It follows now readily from the same arguments that we have developed to show Lemma 2.2 that $\Delta_n^{(\alpha)}$ and $\Delta_{0,n}^{(\alpha)}$ have the same asymptotic behavior. Specifically, we have:

Corollary 2.2. *As $n \to \infty$, there is the weak convergence*

$$\frac{\Delta_n^{(\alpha)}}{\ln^3 n} - 3\frac{\alpha}{1-\alpha} c \ln\ln n \Rightarrow \frac{\alpha}{1-\alpha} c \left(\mathcal{Z}'_{c,\alpha} + 1 - \frac{1}{\alpha} \right)$$

where $\mathcal{Z}'_{c,\alpha}$ is the random variable defined in (2.26).

Similarly as in Section 2.2.2, we now resume the growth of the system of branching process with rare mutation from the size $(2+a)\lfloor \ln^4 n \rfloor + a$ to the size $(2+a)n + a$, and show that the fluctuations of $\Delta_n^{(\alpha)}$ spread regularly. In this direction, we write $Z'^{(a)} = (Z'^{(a)}(t) : t \geq 0)$ for the process of the total size of the population started from $Z'^{(a)}(0) = (2+a)\lfloor \ln^4 n \rfloor + a$, which has the same law as the branching process $Z^{(a)}$. We observe that the population at the time when we restart our observation consists of $\Delta_n^{(\alpha)}$ mutants and $(2+a)\lfloor \ln^4 n \rfloor + a - \Delta_n^{(\alpha)}$ individuals with the ancestral type. Then, we write $Z_0'^{(p,a)} = (Z_0'^{(p,a)}(t) : t \geq 0)$ for the process that counts the number of clone individuals, which has the same law as $Z_0^{(p,a)}$ but starting from $Z_0'^{(p,a)}(0) = (2+a)\lfloor \ln^4 n \rfloor + a - \Delta_n^{(\alpha)}$. We consider the time

$$\tau'^{(a)}(n) = \inf\{t \geq 0 : Z'^{(a)}(t) = (2+a)n + a\}.$$

Then the number of individuals with the ancestral type at time when the total population generated by the branching process reaches $(2+a)n + a$ is given by

$$G_n^{(\alpha)} = Z_0'^{(p,a)}(\tau'^{(a)}(n)).$$

We are now able to state the following analog of Theorem 2.2.

Theorem 2.4. *Set $\alpha = (1+a)/(2+a)$. As $n \to \infty$, there is the weak convergence*

$$\left(n^{-1} G_n^{(\alpha)} - \frac{1}{1-\alpha} e^{-\alpha c} \right) \ln n - \frac{\alpha}{1-\alpha} c e^{-\alpha c} \ln\ln n \Rightarrow -\frac{\alpha}{1-\alpha} c e^{-\alpha c} \left(\mathcal{Z}'_{c,\alpha} + 1 - \frac{1}{\alpha} \right),$$

where $\mathcal{Z}'_{c,\alpha}$ is the random variable defined in (2.26).

Proof. We recall that

$$W'^{(a)}(t) := e^{-(2+a)t} Z'^{(a)}(t) \quad \text{and} \quad W_0'^{(p,a)}(t) := e^{-(1+p(1+a))t} Z'^{(p,a)}(t), \qquad t \geq 0$$

are nonnegative square-integrable martingales which converge a.s. and in $L^2(\mathbb{P})$. Hence from the estimate of Equation (6) in [75], we get for all $\eta > 0$ that

$$\lim_{n\to\infty} \mathbb{P}\left(\left| ((2+a)n + a) e^{-(2+a)\tau'^{(a)}(n)} - ((2+a)\lfloor \ln^4 n \rfloor + a) \right| > \eta \ln^3 n \right) = 0,$$

this yields

$$e^{(2+a)\tau'^{(a)}(n)} = \frac{n}{\ln^4 n} + o\left(\frac{1}{\ln n} \right) \qquad \text{in probability.}$$

On the other hand, using the fact that $Z_0'^{(p,a)}(0) \leq (2+a)\lfloor \ln^4 n \rfloor + a$, we also get

$$\lim_{n\to\infty} \mathbb{P}\left(\left| e^{-(1+p(1+a))\tau'^{(a)}(n)} Z_0'^{(p,a)}(\tau'^{(a)}(n)) - Z_0'^{(p,a)}(0) \right| > \eta \ln^3 n \right) = 0,$$

and deduce that

$$G_n^{(\alpha)} = e^{(1+p(1+a))\tau'^{(a)}(n)}((2+a)\ln^4 n - \Delta_n^{(\alpha)}) + o\left(\frac{n}{\ln n}\right) \qquad \text{in probability.}$$

Skorokhod's representation theorem enables us to assume that the weak convergence in Corollary 2.2 holds almost surely. Hence

$$G_n^{(\alpha)} = e^{(1+p(1+a))\tau'^{(a)}(n)} \left((2+a)\ln^4 n - \frac{\alpha}{1-\alpha} c \ln^3 n \left(3\ln\ln n + \mathcal{Z}'_{c,\alpha} + 1 - \frac{1}{\alpha} \right) \right) + o\left(\frac{n}{\ln n}\right)$$

in probability. It follows that

$$G_n^{(\alpha)} = \frac{1}{1-\alpha} e^{-\alpha c} n + \frac{\alpha}{1-\alpha} c e^{-\alpha c} n \frac{\ln\ln n}{\ln n} - \frac{\alpha}{1-\alpha} c e^{-\alpha c} \frac{n}{\ln n} \left(\mathcal{Z}'_{c,\alpha} + 1 - \frac{1}{\alpha} \right) + o\left(\frac{n}{\ln n}\right)$$

in probability, which completes the proof. $\square$

We have now all the ingredients to establish Theorem 2.3.

Proof of Theorem 2.3. We follow Bertoin and Uribe Bravo [32], and we consider a continuous time version of the growth algorithm with preferential attachment as we have done for the b-ary recursive trees. We start at 0 from the tree $\{0,1\}$, and once the random tree with size $n \geq 2$ has been constructed, we equip each vertex $i \in \{0,1,\dots,n\}$ with an exponential random variable ζ_i of parameter $d_n(i)+a$, independently of the other vertices. Then the next vertex $n+1$ is attached after time $\min_{i\in\{0,1,\dots,n\}} \zeta_i$ at the vertex $v_n = \mathrm{argmin}_{i\in\{0,1,\dots,n\}} \zeta_i$. Let us denote by $T^{(a)}(t)$ the tree which has been constructed at time t, and by $|T^{(a)}(t)|$ its size, i.e. its number of vertices. It should be plain that if we define

$$\gamma^{(a)}(n) = \inf\{t \geq 0 : |T^{(a)}(t)| = n+1\},$$

then $T^{(a)}\left(\gamma^{(a)}(n)\right)$ is a version of a scale-free tree of size $n+1$, $T_n^{(a)}$. Furthermore, the process $Y^{(a)}$ defined by

$$Y^{(a)}(t) = (2+a)|T^{(a)}(t)| - 2, \qquad t \geq 0,$$

is a pure branching process with initial value $Y^{(a)}(0) = 2a+2$ that has only jumps of size $2+a$, and with unit birth rate per unit population size. Then we incorporate Bernoulli bond percolation to the algorithm similarly to how we did in Section 2.2 for the b-ary recursive trees. We draw an independent Bernoulli random variable ϵ_p with parameter p, each time an edge is inserted. If $\epsilon_p = 1$, the edge is left intact, otherwise we cut this edge at its midpoint. We write $T^{(p,a)}(t)$ for the resulting combinatorial structure at time t. Hence the percolation clusters of $T^{(a)}(t)$ are the connected components by a path of intact edges of $T^{(p,a)}(t)$. Let $T_0^{(p,a)}(t)$ be the subtree that contains the root. We write $H_0^{(p,a)}(t)$ for the number of half-edges pertaining to the root cluster at time t and set

$$Y_0^{(p,a)}(t) = (2+a)|T_0^{(p,a)}(t)| + H_0^{(p,a)}(t) - 2.$$

We now observe the connection with the system of branching processes with rare mutations $\mathbf{Z}^{(p,a)}$. It should be plain from the construction that $Y_0^{(p,a)}$ has the same random evolution as the process of the number of individuals with the ancestral type $Z_0^{(p,a)}(t)$. In fact, the process $Y^{(a)}$ coincides with the process of the total size $Z^{(a)}$. Then, the size $\Gamma_n^{(\alpha)}$ of the percolation cluster containing the root when the structure has size $n+1$ satisfies $\Gamma_n^{(\alpha)} = |T_0^{(p,a)}(\gamma^{(a)}(n))|$. In addition, it should be plain that

$$Y_0^{(p,a)}(\gamma^{(a)}(n)) = (2+a)\Gamma_n^{(\alpha)} + H_0^{(p,a)}(\gamma^{(a)}(n)) - 2, \tag{2.30}$$

coincides with the number of individuals with the ancestral type in the branching system $\mathbf{Z}^{(p,a)}$, at time when the total population reaches the size $(2+a)n+a$, i.e. $G_n^{(\alpha)}$. Then, in order to establish Theorem 2.3, it is sufficient to get an estimate of the number of half-edges pertaining to the root-subtree at time $\gamma^{(a)}(n)$. We follow the route of Lemma 2.8 and observe that the process

$$L^{(p,a)}(t) := H_0^{(p,a)}(t) - \frac{1-p}{1+p+pa} Y_0^{(p,a)}(t), \qquad t \geq 0$$

is a martingale whose jumps have size at most $2+a$. Since there are at most n jumps up to time $\gamma^{(a)}(n)$, the bracket of $L^{(p,a)}$ can be bounded by $[L^{(p,a)}]_{\gamma^{(a)}(n)} \leq (2+a)^2 n$. Hence we have

$$\lim_{n\to\infty} \mathbb{E}\left(\left| \frac{\ln n}{n} L^{(p,a)}(\gamma^{(a)}(n)) \right|^2 \right) = 0.$$

On the other hand, from Lemma 3 in [32] we get that

$$\lim_{n\to\infty} e^{-(2+a)\gamma^{(a)}(n)} Y^{(a)}(\gamma^{(a)}(n)) = \lim_{n\to\infty} e^{-(1+p(1+a))\gamma^{(a)}(n)} Y_0^{(p,a)}(\gamma^{(a)}(n)) = W^{(a)}(\infty)$$

in probability, where $W^{(a)}(\infty)$ is defined as the terminal value of the martingale $W^{(a)}(t) = e^{-(2+a)t} Y^{(a)}(t)$. Thus, we have that

$$\lim_{n\to\infty} \frac{\ln n}{n} H_0^{(p,a)}(\gamma^{(a)}(n)) = c e^{-\alpha c} \qquad \text{in probability,}$$

and the result in Theorem 2.3 follows from Theorem 2.4 and the identity (2.30). $\square$

CHAPTER 3

Cutting-down random trees

"It's the questions we can't answer that teach us the most. They teach us how to think. If you give a man an answer, all he gains is a little fact. But give him a question and he'll look for his own answers."

— Patrick Rothfuss, The Wise Man's Fear

In this chapter, we study the asymptotic behavior of the cut-tree associate with the destruction procedure of a random tree described in the Section 1.3. For instance, we consider uniform random recursive trees, binary search trees and scale-free random trees. This is based on the article [3].

3.1 Introduction and main result

3.1.1 General introduction

Let us first recall the definition of cut-tree introduced by Bertoin [39]. Consider a tree T_n on a finite set of vertices, say $[n] := \{1, \dots, n\}$, rooted at 1. We associate with T_n a random rooted binary tree Cut(T_n) with n leaves which records the genealogy induced by the the destruction process of T_n: each vertex of Cut(T_n) corresponds to a subset (or block) of $[n]$, the root of Cut(T_n) is the entire set $[n]$ and its leaves are the singletons $\{1\}, \dots, \{n\}$. We remove successively the edges of T_n in a uniform random order; at each step, a subtree of T_n with set of vertices, say, B , splits into two subtrees with sets of vertices, say, B' and B'' respectively; in Cut(T_n), B' and B'' are the two offsprings of B . Notice that, by construction, the set of leaves of the subtree of Cut(T_n) that stems from some block coincides with this block. See Figure 3.1 for an illustration.

The cut-tree is an interesting tool to study the isolation of nodes in a tree. As an example, sample k vertices uniformly at random with replacement in the tree T_n, say, $U_1, \dots, U_k$. Then independently, start the fragmentation of the tree as described above but in addition, each time the removal of an edge creates a connected component which does not contain any of the k selected vertices, discard immediately the latter component. This dynamics stop when the tree is reduced to the k selected singletons. It should be plain that the subtrees which are not discarded correspond to the blocks of the tree Cut(T_n) reduced to its branches from its root to the k leaves $\{U_1\}, \dots, \{U_k\}$. As a consequence, the number of steps of this isolation procedure is given by the number of internal nodes of this reduced tree or, equivalently, its length minus the number of distinct leaves plus one; in the case $k = 1$, the latter is simply the height

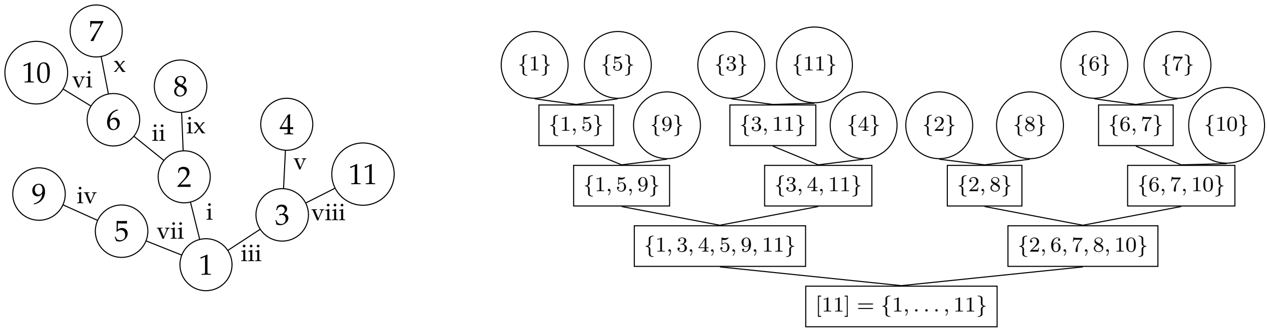

FIGURE 3.1: A tree of size eleven with the order of cuts on the left, and the corresponding cut-tree on the right

of a uniform random leaf in $\text{Cut}(T_n)$.

On the other hand, Baur [76] has recently introduced another tree associated to the destruction process of uniform random recursive trees, called *tree of components*. Informally, one considers a dynamically version of the cutting procedure, where edges are equipped with i.i.d. exponential clocks and deleted at time given by the corresponding variable. Then, each removal of an edge gives birth to a new tree component, which sizes and birth times are encoding by a tree-indexed process. He used this tree of components to study cluster sizes created from performing Bernoulli bond percolation on uniform random recursive trees. We do not study the tree of components in this work but, we think would be of interest, and may be seen as a complement of the cut-tree. However, a common feature with our analysis is that, it is useful to consider a continuous time version of the destruction process.

In this chapter, we study the behavior of $\text{Cut}(T_n)$ when the tree T_n is star-shape, that is, the last common ancestor of two randomly chosen vertices is close to the root (after proper rescaling) with high probability. We consider also that T_n has a small height of order $o(\sqrt{n})$, in the sense that that the distance (the number of edges) between its root 1, and a typical vertex in T_n is of this order $o(\sqrt{n})$ For instance, this is the case for uniform random recursive trees, binary search trees, scale-free random trees and regular trees; see for example Drmota [8] and Barabási [11]. Informally, the main result provides a general criterion, depending on the nature of T_n, for the convergence in distribution of the rescaled $\text{Cut}(T_n)$ when $n \to \infty$.

We next introduce the necessary notation and recall some relevant results on the topological space of trees. This, we will enables us to state our main result in Section 3.1.3.

3.1.2 Measured metric spaces and the Gromov-Prokhorov topology

We begin by introducing some basic facts about topological space of trees in which limits can be taken, and define the limit objects. A pointed metric measure space is a quadruple $(\mathcal{T}, d, \rho, \nu)$ where $(\mathcal{T}, d)$ is a separable and complete metric space, $\rho \in \mathcal{T}$ a distinguished element called the root of $\mathcal{T}$, and ν a Borel probability measure on $(\mathcal{T}, d)$. This quadruple is called a real tree if in addition, $\mathcal{T}$ is a tree, in the sense that it is a geodesic space for which any two points are connected via a unique continuous injective path up to re-parametrization. This is a continuous analog of the graph-theoretic definition of a tree as a connected graph with no cycle. For sake of simplicity, we frequently write $\mathcal{T}$ to refer to a pointed metric measure space $(\mathcal{T}, d, \rho, \nu)$. We say that two measured rooted spaces $(\mathcal{T}, d, \rho, \nu)$ and $(\mathcal{T}', d', \rho', \nu')$ are isometry-equivalent if there exists a root-preserving, bijective isometry $\phi : \text{supp}(\mu) \cup \{\rho\} \to \mathcal{T}'$ (here supp is the topological support) such that the image of ν by ϕ is ν'. This defines an equivalence relation between pointed metric measure spaces, and we note that representatives $(\mathcal{T}, d, \rho, \nu)$ of a given isometry-equivalence class can always be assumed to have $\text{supp}(\mu) \cup \{\rho\} = \mathcal{T}$. It is also convenient to agree that for $a > 0$, $a\mathcal{T}$ denotes the same space $\mathcal{T}$ but with distance rescaled by the factor a, i.e. $(\mathcal{T}, ad, \rho, \nu)$.

It is well-known that the set $\mathbb{M}$ of isometry-equivalence classes of pointed metric spaces is a Polish space when endowed with the so-called Gromov-Prokhorov topology see Section 1.3.2. This topology was introduced by Greven, Pfaffelhuber and Winter in [77] under the name of *Gromov-weak* topology. We can then view the $\text{Cut}(T_n)$ for $n \geq 1$ as a sequence random variables with values in $\mathbb{M}$ (i.e. a sequence of real random tree). For convenience, we adopt a slightly different point of view for $\text{Cut}(T_n)$ than the usual for finite trees, focusing on leaves rather than internal nodes. More precisely, we set $[n]^0 = \{0, 1, \dots, n\}$ where 0 correspond to the root $[n]$ of $\text{Cut}(T_n)$ and $1, \dots, n$ to the leaves (i.e. i is identified with the singleton $\{i\}$). We consider the random pointed metric measure space $([n]^0, \delta_n, 0, \mu_n)$ where δ_n is the random graph distance on $[n]^0$ induced by the cut-tree, 0 is the distinguished element, and μ_n is the uniform probability measure on $[n]$ extended by $\mu_n(0) = 0$. That is, μ_n is the uniform probability measure on the set of leaves of $\text{Cut}(T_n)$. We point out that the combinatorial structure of the cut-tree can be recovered from $([n]^0, \delta_n, 0, \mu_n)$, so by a slight abuse of notation, sometimes we refer to $\text{Cut}(T_n)$ as the latter pointed metric measure space.

Finally, we recall a convenient characterization of the Gromov-Prokhorov topology that relies on the convergence of distances between random points. A sequence $(\mathcal{T}_n, d_n, \rho_n, \nu_n)$ of pointed measure metric spaces converges in the Gromov-Prokhorov sense to an element of $\mathbb{M}$, say $(\mathcal{T}_\infty, d_\infty, \rho_\infty, \nu_\infty)$, if and only if the following holds: for $n \in \{1, 2, \dots\} \cup \{\infty\}$, set $\xi_n(0) = \rho_n$ and let $\xi_n(1), \xi_n(2), \dots$ be a sequence of i.i.d. random variables with law ν_n, then

$$(d_n(\xi_n(i), \xi_n(j)) : i, j \geq 0) \Rightarrow (d_\infty(\xi_\infty(i), \xi_\infty(j)) : i, j \geq 0)$$

where $\Rightarrow$ means convergence in the sense of finite-dimensional distribution, $\xi_\infty(0) = \rho_\infty$ and $\xi_\infty(1), \xi_\infty(2), \dots$ is a sequence of i.i.d. random variables with law ν_∞; see for example Corollary 8 of [49]. One can interpret $(d_\infty(\xi_\infty(i), \xi_\infty(j)) : i, j \geq 0)$ as the matrix of mutual distances between the points of an i.i.d. sample of $(\mathcal{T}_\infty, d_\infty, \rho_\infty, \nu_\infty)$. Moreover, it is important to point out that by the Gromov's reconstruction theorem in [50], the distribution of the above matrix of distances characterizes $(\mathcal{T}_\infty, d_\infty, \rho_\infty, \nu_\infty)$ as an element of $\mathbb{M}$.

3.1.3 Main result

We first introduce notation and hypotheses which will have an important role for the rest of the work. Recall that T_n is a tree with set of vertices $[n] = \{1, \dots, n\}$, rooted at 1. We denote by u and v two independent uniformly distributed random vertices on $[n]$. Let d_n be the graph distance in T_n, and

$\ell : \mathbb{N} \to \mathbb{R}_+$ be some function such that $\lim_{n\to\infty} \ell(n) = \infty$. We introduce the following hypothesis

$$\frac{1}{\ell(n)}(d_n(1,u), d_n(u,v)) \Rightarrow (\zeta_1, \zeta_1 + \zeta_2). \tag{H}$$

where ζ_1 and ζ_2 are i.i.d. variables in $\mathbb{R}_+$ with no atom at 0. This happens with ζ_i a positive constant for some important families of random trees, such as uniform recursive trees, regular trees, scale-free random trees and binary search trees (and more generally b-ary recursive trees). In Section 3.4, we consider a different class of examples where the variable ζ_i is not a constant, which results of the mixture of similar trees satisfying the hypothesis (H).

Remark 3.1. *We observe that*

$$d_n(u,v) = d_n(1,u) + d_n(1,v) - 2d_n(1, u\wedge v),$$

where $u \wedge v$ is the last common ancestor of u and v in T_n. Then, the condition (H) readily implies that

$$\lim_{n\to\infty} \ell(n)^{-1} d_n(1, u\wedge v) = 0,$$

in probability. Moreover, if for each fixed $k \in \mathbb{N}$, we denote by $L_{k,n}$ the length of the tree T_n reduced to k vertices chosen uniformly at random with replacement and its root 1, i.e. the minimal number of edges of T_n which are needed to connect 1 and such vertices, we see that (H) is equivalent to

$$\frac{1}{\ell(n)}(L_{1,n}, L_{2,n}) \Rightarrow (\zeta_1, \zeta_1 + \zeta_2).$$

We then write

$$\lambda(t) = \mathbb{E}[e^{-t\zeta_1}], \qquad \text{for} \ \ t \geq 0,$$

for the Laplace transform of the random variable ζ_1. We henceforth denote

$$a = \mathbb{E}[1/\zeta_1],$$

which can be infinite. We define the bijective mapping $\Lambda : [0,\infty) \to [0,a)$ by

$$\Lambda(t) = \int_0^t \lambda(s) \mathrm{d}s, \qquad \text{for} \ \ t \geq 0,$$

where $\Lambda(\infty) = \lim_{t\to\infty} \Lambda(t) = a$, and write Λ^{-1} for its inverse mapping. Observe that (H) entails that

$$\frac{1}{\ell(n)} d_n(u,v) \Rightarrow \zeta_1 + \zeta_2,$$

then we consider the next technical condition

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{\ell(n)}{d_n(u,v)} \mathbf{1}_{\{u\neq v\}}\right] = \mathbb{E}\left[\frac{1}{\zeta_1+\zeta_2}\right] < \infty. \tag{H'}$$

Theorem 3.1. *Suppose that (H) and (H′) hold with ℓ such that $\ell(n) = o(\sqrt{n})$. Furthermore, assume that $a < \infty$. Then as $n \to \infty$, we have the following convergence in distribution in the sense of the pointed Gromov-Prokhorov topology:*

$$\frac{\ell(n)}{n}\mathrm{Cut}(T_n) \Rightarrow I_\mu.$$

where I_μ is the pointed measure metric space given by the interval $[0, a]$, pointed at 0, equipped with the Euclidean distance, and the probability measure μ given by

$$\int_0^a f(x)\mu(\mathrm{d}x) = -\int_0^a f(x)\,\mathrm{d}\lambda \circ \Lambda^{-1}(x) \tag{3.1}$$

where f is a generic positive measurable function. The result still valid when $a = \infty$, and then one considers the interval $[0, \infty)$, pointed at 0, equipped with the same distance and measure.

Theorem 3.1 does not apply for the family of critical Galton-Watson trees conditioned to have size n considered for Bertoin and Miermont [38] and Dieuleveut [40] since they do not satisfy the condition (H), and the height of a typical vertex is not of the order $o(\sqrt{n})$. For instance, the case when T_n is a Cayley tree (conditioned Galton-Watson tree with Poisson offspring distribution), for which it is know that $\ell(n) = \sqrt{n}$ and the variable $L_{i,n}$ in Remark 3.1, for $i = 1, 2$, is a chi-variable with $2k$ degrees of freedom; see for example Aldous [61]. Then, we believe that the threshold $\sqrt{n}$ appearing in this work is critical, i.e. for trees with larger heights the limit of their rescaled cut-tree is a random tree, and not a deterministic tree. For instance, when $T_n^{(c)}$ is a Cayley tree of size n, it has been shown in [37] that $n^{-1/2}\mathrm{Cut}(T_n^{(c)})$ converges in distribution to a Brownian Continuum Random tree, in the sense of Gromov-Hausdorff-Prokhorov. This has been extended in [38] to a large family of critical Galton-Watson trees with finite variance, and by Dieuleveut [40] when the offspring distribution belongs to the domain of attraction of a stable law of index $\alpha \in (1, 2]$, both in the sense of Gromov-Prokhorov. We point out that in [40] the limit is a α-stable continuum random.

On the other hand, Bertoin [39] has shown that for uniform random recursive tree $T_n^{(r)}$ of size n that upon rescaling the graph distance of $\mathrm{Cut}(T_n^{(r)})$ by a factor $n^{-1}\ln n$, the latter converges in probability in the sense of pointed Gromov-Hausdorff-Prokhorov distance to the unit interval $[0, 1]$ equipped with the Euclidean distance and the Lebesgue measure, pointed at 0. The proof of this result uses crucially a coupling due to Iksanov and Möhle [26] that connects the destruction process in this family of trees with a remarkable random walk, which is not fulfilled in general for trees we are interested. Thus, we have to used a fairly different route.

The plan of the rest of this chapter is as follows. In section 3.2, we introduce a continuous version of the cutting down procedure, where edges are equipped with i.i.d. exponential random variables and removed at time given by the corresponding variable. Following Bertoin [52] we then represent the destruction process up to a certain finite time as a Bernoulli bond-percolation, allowing us relate the tree components with percolation clusters. We then study the behavior of the number of edges which are remove from the root cluster as time passed, which is closely related with the distance induced by the cut-tree. We then establish our main result Theorem 3.1 in Section 3.3. In Section 3.4, we provide some examples of trees that fulfill the hypotheses (H) and (H'). Then in Section 3.5 we present some applications on the isolation of multiple vertices, which extend the results of Kuba and Panholzer [53], and Baur and Bertoin [54] for uniform random recursive trees. Finally, Section 3.6 is devoted to the proof of a technical result about the shape of scale-free random trees, which may be of independent interest.

3.2 Cutting down in continuous time

The purpose of this section is to study the destruction dynamics on a general sequence of random trees T_n. We consider a continuous time version of the destruction process in which edges are removed independently one of the others at a given rate. We establish the link with Bernoulli bond-percolation and deduce some properties related to the destruction process, which will be relevant for the proof of Theorem 3.1.

Recall that for each fixed $k \in \mathbb{N}$, we denote by $L_{k,n}$ the length of the tree T_n reduced to k vertices chosen uniformly at random with replacement and its root 1. Recall also the Remark 3.1 and then consider the following weaker version of the hypothesis (H),

$$\frac{1}{\ell(n)} L_{k,n} \Rightarrow \zeta_1 + \cdots + \zeta_k, \qquad (H_k)$$

where $\zeta_1, \ldots$ is a sequence of i.i.d. variables in $\mathbb{R}_+$ with no atom at 0, and the convergence in (H_k) is in the sense of one-dimensional distribution, i.e. for each fixed k. We stress that the hypothesis (H) implies (H_k) for $k = 1, 2$.

We then present the continuous time version of the destruction process. We attach to each edge e of T_n an independent exponential random variable $\mathbf{e}(e)$ of parameter $1/\ell(n)$, and we delete it at time $\mathbf{e}(e)$. After the $(n-1)$th edge has been deleted, the tree has been destructed, and the process ends. Rigorously, let $e_1, \ldots, e_{n-1}$ denote the edges of T_n listed in the increasing order of their attached exponential random variables, i.e. such that $\mathbf{e}(e_1) < \cdots < \mathbf{e}(e_{n-1})$. Then at time $\mathbf{e}(e_1)$, the first edge e_1 is removed from T_n, and T_n splits into two subtrees, say τ_n^1 and τ_n^*, where τ_n^1 contains the root 1. Next, if e_2 connects two vertices in τ_n^* then at time $\mathbf{e}(e_2)$, τ_n^* splits in two tree components. Otherwise, τ_n^1 splits in two subtrees after removing the edge e_2. We iterate in an obvious way until all the vertices of T_n have been isolated.

Define $p_n(t) = \exp(-t/\ell(n))$ for $t \geq 0$, and observe that the probability that a given edge has not yet been removed at time t in the continuous time destruction process is $p_n(t)$. Thus, the configuration observed at time t is precisely that resulting from a Bernoulli bond percolation on T_n with parameter $p_n(t)$. Further, Bertoin [18] proved that when the hypothesis (H_k) is fulfilled for $k = 1, 2$, the percolation parameter $p_n(t)$ corresponds to the supercritical regime, in the sense that with high probability, there exists a giant cluster, that is of size (number of vertices) comparable to that of the entire tree. Thus focusing on the evolution of the tree component which contains the root 1, we write $X_n(t)$ for its size at time $t \geq 0$; plainly $X_n(t) \leq n$. We shall establish the following limit theorem which is an improvement of Corollary 1 (i) in [18].

Proposition 3.1. *Suppose that* (H_k) *holds for* $k = 1, 2$. *Then, we have that*

$$\lim_{n\to\infty} \sup_{s\geq 0} |n^{-1}X_n(s) - \lambda(s)| = 0 \qquad \textit{in probability}. \tag{3.2}$$

Proof. It follows from Corollary 1(i) in [18] that for $t \geq 0$

$$\lim_{n\to\infty} n^{-1}X_n(t) = \lambda(t) \qquad \text{in probability},$$

where $\lambda(t) = \mathbb{E}(e^{-t\zeta_1})$ for $t \geq 0$, when ever (H_k) holds for $k = 1, 2$. Then by the diagonal procedure, we may extract from an arbitrary increasing sequence of integers a subsequence, say $(n_l)_{l\in\mathbb{N}}$, such that with probability one,

$$\lim_{l\to\infty} n_l^{-1}X_{n_l}(s) = \lambda(s) \qquad \text{for all rational} \;\; s \geq 0.$$

As $s \to X_n(s)$ decreases, and $s \to \lambda(s)$ is continuous, the above convergence holds uniformly on $[0, t]$ for an arbitrary fixed $t > 0$, i.e.

$$\lim_{l\to\infty} \sup_{0\leq s\leq t} |n_l^{-1}X_{n_l}(s) - \lambda(s)| = 0 \qquad \text{a.s.}. \tag{3.3}$$

On the other hand, we observe that $\lim_{s\to\infty}\lambda(s) = 0$. Then for $\varepsilon > 0$, we can find $t_\varepsilon > 0$ and $N(\varepsilon) > 0$ such that

$$\sup_{s>t_\varepsilon} |n_l^{-1}X_{n_l}(s) - \lambda(s)| < \varepsilon \qquad \text{for} \;\; n_l > N(\varepsilon), \quad \text{a.s.},$$

and therefore, our claim follows by combining (3.3) and the above observation. □

It is interesting to recall that the reciprocal of Proposition 3.1 holds. More precisely, Corollary 1 (ii) in [18] shows that (H_k), for $k = 1, 2$, form a necessarily and sufficient condition for (3.2).

In order to make the connexion with the discrete destruction process introduced at the beginning of this work, which is the one we are interested in, we now turn our attention to the number $R_n(t)$ of edges of the current root component which have been removed up to time t in the procedure described above. We observe that every jump of the process $R_n = (R_n(t) : t \geq 0)$ corresponds to removing an edge from the root component according to the discrete destruction process. We interpret the latter as a continuous time version of a random algorithm introduced by Meir and Moon [24, 33] for the isolation of the root. Recall also that

$$\Lambda(t) = \int_0^t \lambda(s)\mathrm{d}s, \qquad \text{for} \;\; t \geq 0.$$

Lemma 3.1. *Suppose that* (H_k) *holds for* $k = 1, 2$, *with* ℓ *such that* $\ell(n) = o(\sqrt{n})$. *Then, we have for every fixed* $t > 0$

$$\lim_{n\to\infty} \sup_{0\leq s\leq t} \left| \frac{\ell(n)}{n} R_n(s) - \Lambda(s) \right| = 0 \qquad \textit{in probability}.$$

Proof. We denote by $X_n = (X_n(t) : t \geq 0)$ the process of the size of the root cluster. The dynamics of the continuous time destruction process show that the counting process R_n grows at rate $\ell(n)^{-1}(X_n - 1)$, which means rigorously that the predictable compensator of $R_n(t)$ is absolutely continuous with respect to the Lebesgue measure with density $\ell(n)^{-1}(X_n(t) - 1)$. In other words,

$$M_n(t) = R_n(t) - \int_0^t \ell(n)^{-1}(X_n(s) - 1)\mathrm{d}s$$

is a martingale; note also that its jumps $|M_n(t) - M_n(t-)|$ have size at most 1. Since there are at most $n-1$ jumps up to time t, the bracket of M_n can be bounded by $[M_n]_t \leq n - 1$. By Burkholder–Davis–Gundy inequality, we have that

$$\mathbb{E}[|M_n(t)|^2] \leq n - 1,$$

and in particular, since we assumed that $\ell(n) = o(\sqrt{n})$,

$$\lim_{n\to\infty} \mathbb{E}\left[\left|\frac{\ell(n)}{n} M_n(t)\right|^2\right] = 0. \tag{3.4}$$

On the other hand, since (H_k) holds for $k = 1, 2$, Proposition 3.1 and dominated convergence entail

$$\lim_{n\to\infty} \frac{\ell(n)}{n} \int_0^t \ell(n)^{-1}(X_n(s) - 1)\mathrm{d}s = \int_0^t \lambda(s)\mathrm{d}s \quad \text{in probability.}$$

Hence from (3.4) we have that

$$\lim_{n\to\infty} \frac{\ell(n)}{n} R_n(t) = \Lambda(t) \quad \text{in probability,}$$

and since $t \to R_n(t)$ increases, by the diagonal procedure as in the proof of Proposition 3.1, our claim follows. □

We continue our analysis of the destruction process, and prepare the ground for the main result of this section, which is the estimation of the number of steps in the algorithm for the isolating the root which are needed to disconnect (and not necessarily isolate) a vertex chosen uniformly at random from the root component. We start by studying the analogous quantity in continuous time. For each fixed $n \in \mathbb{N}$, we denote by $u_1, u_2, \ldots$ a sequence of i.i.d. vertices in $[n] = \{1, \ldots, n\}$ with the uniform distribution. Next, for every $i \in \mathbb{N}$, we write $\Gamma_i^{(n)}$ the first instant when the vertex u_i is disconnected from the root component. We shall establish the following limit theorem in law.

Proposition 3.2. *Suppose that (H_k) holds for $k = 1, 2$. Then as $n \to \infty$, the random vector*

$$(\Gamma_i^{(n)} : i \geq 1) \Rightarrow (\gamma_i : i \geq 1)$$

in the sense of finite-dimensional distribution, where $\gamma_1, \gamma_2, \ldots$ are i.i.d. random variables in $\mathbb{R}_+$ with distribution given by $\mathbb{P}(\gamma_1 > t) = \lambda(t)$ for $t \geq 0$.

Proof. We observe that for every $j \in \mathbb{N}$ and $t_1, \ldots, t_j \geq 0$, there is the identity

$$\mathbb{P}(\Gamma_1^{(n)} > t_1, \ldots, \Gamma_j^{(n)} > t_j) = \mathbb{P}(u_1 \in T_n^{(1)}(t_1), \ldots, u_j \in T_n^{(1)}(t_j)),$$

where $T_n^{(1)}(t)$ denotes the subtree at time t which contains the root 1. Recall that $u_1, \ldots, u_j$ are i.i.d. uniformly distributed vertices, which are independent of the destruction process. On the other hand, for $t \geq 0$ the variable $n^{-1}X_n(t)$ is the proportion of vertices in the root component at time t, and represents the conditional probability that a vertex of T_n chosen uniformly at random belongs to the root component at time t. We thus have

$$\mathbb{P}(\Gamma_1^{(n)} > t_1, \ldots, \Gamma_j^{(n)} > t_j) = \mathbb{E}\left[n^{-j}\prod_{i=1}^{j} X_n(t_i)\right].$$

Since (H_k) holds for $k = 1, 2$, we conclude from Proposition 3.1 that

$$\lim_{n\to\infty} \mathbb{P}(\Gamma_1^{(n)} > t_1, \ldots, \Gamma_j^{(n)} > t_j) = \prod_{i=1}^{j} \lambda(t_i),$$

which establishes our claim. □

We are now in position to state the main result of this section. We provide a non-trivial limit in distribution for the number $Y_i^{(n)}$ of cuts (in the algorithm for isolating the root) which are needed to disconnect a vertex chosen uniformly at random, say u_i, from the root component.

Corollary 3.1. *Suppose that (H_k) holds for $k = 1, 2$, with ℓ such that $\ell(n) = o(\sqrt{n})$. Then as $n \to \infty$, we have that*

$$\left(\frac{\ell(n)}{n} Y_i^{(n)} : i \geq 1\right) \Rightarrow (Y_i : i \geq 1)$$

in the sense of finite-dimensional distribution, where $Y_1, Y_2, \ldots$ are i.i.d. random variables on $[0, a)$ where $a = \Lambda(\infty)$, and with distribution given by

$$\mathbb{E}[f(Y_1)] = -\int_0^a f(x) \mathrm{d}\lambda \circ \Lambda^{-1}(x), \tag{3.5}$$

where f is a generic positive measurable function.

Proof. Recall that $R_n(t)$ denotes the number of edges of the root component which have been removed up to time t in the continuous procedure described above. We recall also that $\Gamma_i^{(n)}$ denotes the first instant when the vertex u_i, chosen uniformly at random, has been disconnected from the root component. Hence we have the following identity,

$$Y_i^{(n)} = R_n(\Gamma_i^{(n)}) \quad \text{for } \ i \in \mathbb{N}.$$

It follows from Lemma 3.1 and Proposition 4.3 that

$$\lim_{n\to\infty} \left(\frac{\ell(n)}{n} R_n(\Gamma_i^{(n)}) - \Lambda(\Gamma_i^{(n)})\right) = 0 \quad \text{in probability},$$

and therefore, as $n \to \infty$, we have that

$$\left(\frac{\ell(n)}{n} Y_i^{(n)} : i \geq 1\right) \Rightarrow (\Lambda(\gamma_i) : i \geq 1)$$

in the sense of finite-dimensional distribution, where $\gamma_1, \gamma_2, \ldots$ are i.i.d. random variables in $\mathbb{R}_+$ with distribution given by $\mathbb{P}(\gamma_1 > t) = \lambda(t)$. Finally, we only need to verify that the law of $\Lambda(\gamma_1)$ is given by (3.5). We observe that by dominated convergence λ is differentiable, and we denote by λ' its derivative. Then for f a generic positive measurable function that

$$\mathbb{E}\left[f\left(\Lambda(\gamma_1)\right)\right] = -\int_0^\infty f\left(\Lambda(x)\right)\lambda'(x)\mathrm{d}x.$$

On the other hand, we observe that Λ is an increasing continuous and differentiable function whose derivative is never 0. Hence

$$\begin{aligned}\mathbb{E}\left[f\left(\Lambda(\gamma_1)\right)\right] &= -\int_0^{\Lambda(\infty)} f(x)\,\frac{\lambda' \circ \Lambda^{-1}(x)}{\lambda \circ \Lambda^{-1}(x)}\mathrm{d}x \\ &= -\int_0^{\Lambda(\infty)} f(x)\,\mathrm{d}\lambda \circ \Lambda^{-1}(x),\end{aligned}$$

which completes the proof. □

Corollary 3.1 will have a crucial role in the proof of Theorem 3.1. This result will enable us to get a precise estimate of distances in the cut-tree.

Finally, let $N^{(u)}(n)$ be the number of remaining cuts that are needed to isolate a vertex chosen uniformly at random, say u, once it has been disconnected from the root component. The next proposition establishes a criterion which ensures that $N^{(u)}(n)$ is small compared to $n/\ell(n)$ with high probability. This technical ingredient will be useful later on in the proof of Theorem 3.1.

Proposition 3.3. *Assume that (H) and (H′) hold with ℓ such that $\ell(n) = o(\sqrt{n})$. Then we have*

$$\lim_{n\to\infty} \frac{\ell(n)}{n} N^{(u)}(n) = 0 \quad \textit{in probability.}$$

Proof. We write $R_n^{(u)}(t)$ for the number of edges that have been removed up to time t from the tree component containing the vertex u, and Γ_n the first instant when the vertex u has been disconnected from the root cluster; in particular,

$$\lim_{t\to\infty} R_n^{(u)}(\Gamma_n + t) - R_n^{(u)}(\Gamma_n) = N^{(u)}(n).$$

Let $X_n^{(u)}(t)$ be the size of the subtree containing the vertex u at time t. Since each edge is removed with rate $\ell(n)^{-1}$, independently of the other edges, the process

$$M_n^{(u)}(t) = R_n^{(u)}(\Gamma_n + t) - R_n^{(u)}(\Gamma_n) - \int_0^t \ell(n)^{-1}(X_n^{(u)}(\Gamma_n + s) - 1)\mathrm{d}s, \qquad t \geq 0,$$

is a purely discontinuous martingale with terminal value

$$\lim_{t\to\infty} M_n^{(u)}(t) = N^{(u)}(n) - \int_0^\infty \ell(n)^{-1}(X_n^{(u)}(\Gamma_n + s) - 1)\mathrm{d}s.$$

Further, its bracket can be bounded by $[M_n^{(u)}]_t \leq n-1$. Then since we assume that $\ell(n) = o(\sqrt{n})$,

$$\lim_{n\to\infty} \mathbb{E}\left[\left|\frac{\ell(n)}{n}N^{(u)}(n) - \int_0^\infty n^{-1}(X_n^{(u)}(\Gamma_n + s) - 1)\mathrm{d}s\right|^2\right] = 0.$$

Therefore, it only remains to prove that

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^\infty n^{-1}(X_n^{(u)}(\Gamma_n + s) - 1)\mathrm{d}s\right] = 0. \tag{3.6}$$

Let $T_n^{(u)}(s)$ denote the subtree at time s which contains the vertex u. We observe that

$$\begin{aligned}\mathbb{E}\left[\int_0^\infty n^{-1}(X_n^{(u)}(\Gamma_n + s) - 1)\mathrm{d}s\right] &= \mathbb{E}\left[\int_0^\infty n^{-1}(X_n^{(u)}(s) - 1)\mathbf{1}_{\{\Gamma_n \leq s\}}\mathrm{d}s\right] \\ &= \mathbb{E}\left[\int_0^\infty n^{-1}(X_n^{(u)}(s) - 1)\mathbf{1}_{\left\{1\notin T_n^{(u)}(s)\right\}}\mathrm{d}s\right].\end{aligned}$$

We note that a vertex v chosen uniformly at random in $[n]$ and independent of u belong to the same cluster at time t if and only if no edge on the path form u and v has been removed at time t. Recall that the probability that a given edge has not yet been removed at time t is $\exp(-t/\ell(n))$ in the continuous time destruction process. Recall that d_n denotes the graph distance in T_n, and $u \wedge v$ the last common ancestor of u and v. Then, we have that

$$\begin{aligned}\mathbb{E}\left[n^{-1}(X_n^{(u)}(t) - 1)\mathbf{1}_{\{1\notin T^{(u)}(t)\}}\right] &= n^{-1}\mathbb{E}\left[\sum_{i\in[n]\setminus u}\mathbf{1}_{\left\{i\in T_n^{(u)}(t),\, 1\notin T_n^{(u)}(t)\right\}}\right] \\ &= \mathbb{E}\left[\left(e^{-\frac{d_n(u,v)}{\ell(n)}t} - e^{-\frac{L_{2,n}}{\ell(n)}t}\right)\mathbf{1}_{\{v\neq u\}}\right],\end{aligned}$$

where $L_{2,n}$ is the length of the tree T_n reduced to the vertex u, v and its root. Then,

$$\mathbb{E}\left[\int_0^\infty n^{-1}(X_n^{(u)}(\Gamma_n + s) - 1)\mathrm{d}s\right] = \mathbb{E}\left[\left(\frac{\ell(n)}{d_n(u,v)} - \frac{\ell(n)}{L_{2,n}}\right)\mathbf{1}_{\{v\neq u\}}\right]. \tag{3.7}$$

On the other hand, since

$$\frac{\ell(n)}{L_{2,n}}\mathbf{1}_{\{v\neq u\}} \leq \frac{\ell(n)}{d_n(u,v)}\mathbf{1}_{\{v\neq u\}},$$

it is not difficult to see from Remark 3.1 that the assumption (H') implies that

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{\ell(n)}{L_{2,n}}\mathbf{1}_{\{v\neq u\}}\right] = \mathbb{E}\left[\frac{1}{\zeta_1 + \zeta_2}\right] < \infty.$$

Therefore, we get (3.6) by letting $n \to \infty$ in (3.7). □

3.3 Proof of Theorem 3.1

In this section, we prove our main result, Theorem 3.1, on the behavior of the cut-tree of the random tree T_n. We stress that during the proof we consider that the tree T_n is a deterministic tree. This will clearly imply the result for random trees. In this direction, we recall that we view the Cut(T_n) as the pointed metric measure space $([n]^0, \delta_n, 0, \mu_n)$, where 0 corresponds to the root and $1, \dots, n$ to the leaves, δ_n the graph distance induced by the cut-tree, and μ_n the uniform probability measure on $[n]$ with $\mu_n(0) = 0$. We assume that $a = \Lambda(\infty) < \infty$. We then recall that I_μ denotes the pointed measure metric space given by the interval $[0, a]$, pointed at 0, equipped with the Euclidean distance, and the probability measure μ given in (3.1), i.e.

$$\int_0^a f(x)\mu(\mathrm{d}x) = -\int_0^a f(x)\,\mathrm{d}\lambda \circ \Lambda^{-1}(x),$$

where f is a generic positive measurable function. We stress that in the case $a = \infty$ the proof follows along the same lines as that of $a < \infty$. Then, I_μ denotes the pointed measure metric space given by the interval $[0, \infty)$, pointed at 0, equipped with the Euclidean distance and the measure μ.

We recall that to establish weak convergence in the sense induced by the Gromov-Prokhorov topology, we shall prove the convergence in distribution of the rescaled distances of Cut(T_n). Specifically, for every $n \in \mathbb{N}$, set $\xi_n(0) = 0$ and consider a sequence $(\xi_n(i))_{i \geq 1}$ of i.i.d. random variables with law μ_n. We will prove that

$$\left(\frac{\ell(n)}{n}\delta_n(\xi_n(i), \xi_n(j)) : i, j \geq 0\right) \Rightarrow (\delta(\xi(i), \xi(j)) : i, j \geq 0)$$

in the sense of finite-dimensional distribution, where $\xi(0) = 0$ and $(\xi(i))_{i \geq 1}$ is a sequence of i.i.d. random variables on $\mathbb{R}_+$ with law μ. Furthermore, $\delta(\xi(i), \xi(j)) = |\xi(i) - \xi(j)|$ since δ is the Euclidean distance, and in particular, $\delta(0, \xi(i)) = \xi(i)$.

The key idea of the proof relies in the relationship between the distance in Cut(T_n), and the number of cuts needed to disconnect certain number of vertices in T_n. Indeed, the height of the leaf $\{i\}$ in Cut(T_n) is precisely the number of cuts needed to isolate the vertex i in T_n. Therefore, it will be convenient to think in $(\xi_n(i))_{i \geq 1}$ as a sequence of i.i.d. vertices in $[n]$, with the uniform distribution.

Proof of Theorem 3.1. We observe that for $i \geq 1$,

$$\delta_n(\xi_n(0), \xi_n(i)) = \delta_n(0, \xi_n(i))$$

is precisely the number of cuts which are needed to isolate the vertex $\xi_n(i)$. For each $n \in \mathbb{N}$, we denote by $\delta_n^{(1)}(0, \xi_n(i))$ the number of cuts which are needed to disconnect the vertex $\xi_n(i)$ from the root component, and by $\eta(\xi_n(i))$ the remaining number of cuts which are needed to isolate the vertex $\xi_n(i)$ after it has been disconnected. Clearly, we have

$$\delta_n(0, \xi_n(i)) - \delta_n^{(1)}(0, \xi_n(i)) = \eta(\xi_n(i)).$$

Since the condition (H') holds, Proposition 3.3 implies that $\lim_{n\to\infty} n^{-1}\ell(n)\eta(\xi_n(i)) = 0$ in probability for $i \geq 1$. Therefore, the assumption (H) entails according to Corollary 3.1 that

$$\left(\frac{\ell(n)}{n}\delta_n(0,\xi_n(i)) : i \geq 0\right) \Rightarrow (\xi(i) : i \geq 0)$$

in the sense of finite-dimensional distribution. Essentially, we follow the same argument to show that the preceding also holds jointly with

$$\left(\frac{\ell(n)}{n}\delta_n(\xi_n(i),\xi_n(j)) : i,j \geq 1\right) \Rightarrow (\delta(\xi(i),\xi(j)) : i,j \geq 1) \tag{3.8}$$

which is precisely our statement.

In this direction, for $i, j \geq 1$, we denote by $\delta_n^{(2)}(\xi_n(i),\xi_n(j))$ the number of cuts which are needed to isolate the vertices $\xi_n(i)$ and $\xi_n(j)$. We also write $\delta_n^{(3)}(\xi_n(i),\xi_n(j))$ for the number of cuts (in the algorithm for isolating the root) until for the first time, the vertices $\xi_n(i)$ and $\xi_n(j)$ are disconnected. Hence from the description of the cut-tree, it should be plain that

$$\delta_n(\xi_n(i),\xi_n(j)) = (\delta_n^{(2)}(\xi_n(i),\xi_n(j)) + 1) - (\delta_n^{(3)}(\xi_n(i),\xi_n(j)) - 1). \tag{3.9}$$

Next we observe that

$$\delta_n^{(3)}(\xi_n(i),\xi_n(j)) - \min(\delta_n^{(1)}(0,\xi_n(i)),\delta_n^{(1)}(0,\xi_n(j))) \leq \eta(\xi_n(i)) + \eta(\xi_n(j)),$$

and

$$\delta_n^{(2)}(\xi_n(i),\xi_n(j)) - \max(\delta_n^{(1)}(0,\xi_n(i)),\delta_n^{(1)}(0,\xi_n(j))) \leq \eta(\xi_n(i)) + \eta(\xi_n(j)).$$

Since the assumption (H) and (H') hold, it follows from Proposition 3.3 that

$$\lim_{n\to\infty} \frac{\ell(n)}{n}\left(\eta(\xi_n(i)) + \eta(\xi_n(j))\right) = 0 \qquad \text{in probability.}$$

Moreover, Corollary 3.1 implies that

$$\left(\frac{\ell(n)}{n}\delta_n^{(3)}(\xi_n(i),\xi_n(j)) : i,j \geq 1\right) \Rightarrow (\min(\xi(i),\xi(j)) : i,j \geq 1),$$

and

$$\left(\frac{\ell(n)}{n}\delta_n^{(2)}(\xi_n(i),\xi_n(j)) : i,j \geq 1\right) \Rightarrow (\max(\xi(i),\xi(j)) : i,j \geq 1)$$

hold jointly. Therefore, since δ is the Euclidean distance, the convergence in (3.8) follows from the identity (3.9). $\square$

3.4 Examples

In this section, we present some examples of trees that fulfilled the conditions of Theorem 3.1. But first, we observe that when the hypotheses of the latter are satisfied with $\zeta_1 \equiv 1$, the probability measure μ given in (3.1) corresponds to the Lebesgue measure on the unit interval $[0, 1]$. The above follows from the fact that $\lambda(t) = e^{-t}$ for all $t \geq 0$. Then we have the following interesting consequence of Theorem 3.1.

Corollary 3.2. *Suppose that* (H) *and* (H') *hold, with $\zeta_1 \equiv 1$ and ℓ such that $\ell(n) = o(\sqrt{n})$. Then as $n \to \infty$, we have the following convergence in the sense of the pointed Gromov-Prokhorov topology:*

$$\frac{\ell(n)}{n}\mathrm{Cut}(T_n) \Rightarrow I_1.$$

where I_1 is the pointed measure metric space given by the unit interval $[0, 1]$, pointed at 0, equipped with the Euclidean distance and the Lebesgue measure.

A natural example is the class of random trees with logarithmic heights, i.e. which fulfill hypothesis (H) with $\ell(n) = c \ln n$ for some $c > 0$, such as binary search trees, regular trees, uniform random recursive trees, and more generally scale-free random trees. We are now going to prove that (H') is also satisfied for the previous families of trees and therefore their rescaled cut-tree converges in the sense of Gromov-Prokhorov topology to I_1.

1. Binary search trees. A popular family of random trees used in computer science for sorting and searching data is the binary search tree. More precisely, a binary search tree is a binary tree in which each vertex is associated to a key, where the keys are drawn randomly from an ordered set, we say $\{1, \dots, n\}$, until the set is exhausted. The first key is associated to the root. The next key is placed at the left child of the root if it is smaller than the root's key and placed to the right if it is larger. Then one proceeds progressively, inserting key by key. When all the keys are placed one gets a binary tree with n vertices. For further details, see e.g. [8]. Theorem S1 in Devroye [78] shows that the hypothesis (H) holds with $\ell(n) = 2 \ln n$. Hence in order to be in the framework of Corollary 3.2 all that we need is to check that this family of trees fulfills the hypothesis (H'), namely

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{2\ln n}{d_n(u, v)}\mathbf{1}_{\{u\neq v\}}\right] = \frac{1}{2},$$

where u and v are two vertices chosen uniformly at random with replacement from the binary search tree of size n. In this direction, we pick $0 < \varepsilon < (2\ln 2)^{-1}$ and consider the function ϕ_ε given by $\phi_\varepsilon = 0$ on $[0, \varepsilon]$, $\phi_\varepsilon = 1$ on $[2\varepsilon, \infty)$, and ϕ_ε linear on $[\varepsilon, 2\varepsilon]$. We observe that

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{2\ln n}{d_n(u, v)}\phi_\varepsilon\left(\frac{d_n(u, v)}{2\ln n}\right)\mathbf{1}_{\{u\neq v\}}\right] = \frac{1}{2}\phi_\varepsilon\left(\frac{1}{2}\right).$$

Further, we note that $\phi_\varepsilon(1/2) \to 1$ as $\varepsilon \to 0$. Then, it is enough to prove that

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty} \mathbb{E}\left[\left(\frac{2\ln n}{d_n(u, v)} - \frac{2\ln n}{d_n(u, v)}\phi_\varepsilon\left(\frac{d_n(u, v)}{2\ln n}\right)\right)\mathbf{1}_{\{u\neq v\}}\right] = 0, \tag{3.10}$$

in order to show (H'). We write $X^i(n,k)$ for the number of vertices at distance $k \geq 1$ from the vertex i in a binary search tree of size n. Then

$$\begin{aligned}\mathbb{E}\left[\left(\frac{2\ln n}{d_n(u,v)} - \frac{2\ln n}{d_n(u,v)}\phi_\varepsilon\left(\frac{d_n(u,v)}{2\ln n}\right)\right)\mathbf{1}_{\{u\neq v\}}\right] &\leq \mathbb{E}\left[\frac{2\ln n}{d_n(u,v)}\mathbf{1}_{\{d_n(u,v)\leq 2\varepsilon\ln n,\, u\neq v\}}\right] \\ &\leq \frac{2\ln n}{n^2}\sum_{i=1}^{n}\sum_{k=1}^{\lfloor 2\varepsilon \ln n\rfloor}\frac{1}{k}\mathbb{E}[X^i(n,k)].\end{aligned}$$

Since each vertex in a binary search tree has at most two descendants, we observe that $\mathbb{E}[X^i(n,k)] \leq 3\cdot 2^{k-1}$. Then

$$\begin{aligned}\mathbb{E}\left[\left(\frac{2\ln n}{d_n(u,v)} - \frac{2\ln n}{d_n(u,v)}\phi_\varepsilon\left(\frac{d_n(u,v)}{2\ln n}\right)\right)\mathbf{1}_{\{u\neq v\}}\right] &\leq \frac{3\ln n}{n}\sum_{k=1}^{\lfloor 2\varepsilon\ln n\rfloor}\frac{2^k}{k} \\ &\leq \frac{6\ln n}{n}2^{2\varepsilon\ln n},\end{aligned}$$

and therefore we get (3.10) by letting $n \to \infty$ and $\varepsilon \to 0$.

More generally, one can consider a generalization of the binary search trees, namely the b-ary recursive trees and check that these fulfill the conditions of Corollary 3.2 with $\ell(n) = \frac{b}{b-1}\ln n$; we refer to Devroye [29].

2. Scale free random trees. The scale-free random trees form a family of random trees that grow following a preferential attachment algorithm, and are used commonly to model complex real-word networks; see Barabási and Albert [11]. Specifically, fix a parameter $\alpha \in (-1,\infty)$, and start for $n = 1$ from the tree $T_1^{(\alpha)}$ on $\{1,2\}$ which has a single edge connecting 1 and 2. Suppose that $T_n^{(\alpha)}$ has been constructed for some $n \geq 2$, and for every $i \in \{1,\dots,n+1\}$, denote by $g_n(i)$ the degree of the vertex i in $T_n^{(\alpha)}$. Then conditionally given $T_n^{(\alpha)}$, the tree $T_{n+1}^{(\alpha)}$ is built by adding an edge between the new vertex $n+2$ and a vertex v_n in $T_n^{(\alpha)}$ chosen at random according to the law

$$\mathbb{P}\left(v_n = i | T_n^{(\alpha)}\right) = \frac{g_n(i)+\alpha}{2n+\alpha(n+1)}, \qquad i \in \{1,\dots,n+1\}.$$

We observe that when one lets $\alpha \to \infty$ the algorithm yields an uniform recursive tree. It is not difficult to check that the condition (H) in Corollary 3.2 is fulfilled with $\ell(n) = \frac{1+\alpha}{2+\alpha}\ln n$; see for instance [32]. Then, it only remains to check the hypothesis (H'). We only prove the latter when $\alpha = 0$, the general case follows similarly but with longer computations. We then follow the same route as the case of the binary search trees. Pick $\varepsilon > 0$ and consider the same function ϕ_ε that we defined previously. Therefore, it is enough to show that

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\mathbb{E}\left[\left(\frac{\ln n}{2d_n(u,v)} - \frac{\ln n}{2d_n(u,v)}\phi_\varepsilon\left(\frac{2d_n(u,v)}{\ln n}\right)\right)\mathbf{1}_{\{u\neq v\}}\right] = 0,$$

where u and v are two independent uniformly distributed random vertices on $T_n^{(0)}$. We observe that

$$\mathbb{E}\left[\left(\frac{\ln n}{2d_n(u,v)} - \frac{\ln n}{2d_n(u,v)}\phi_\varepsilon\left(\frac{2d_n(u,v)}{\ln n}\right)\right)\mathbf{1}_{\{u\neq v\}}\right] \leq \mathbb{E}\left[\frac{\ln n}{2d_n(u,v)}\mathbf{1}_{\{d_n(u,v)\leq \frac{1}{2}\varepsilon \ln n,\, u\neq v\}}\right]. \tag{3.11}$$

We write $Z^i(n,k)$ for the number of vertices at distance $k \geq 1$ from the vertex i. Then,

$$\begin{aligned}\mathbb{E}\left[\frac{\ln n}{2d_n(u,v)}\mathbf{1}_{\{u\neq v\}}\mathbf{1}_{\{d_n(u,v)\leq \frac{1}{2}\varepsilon \ln n\}}\right] &\leq \frac{\ln n}{n^2}\sum_{i=1}^{n+1}\sum_{k=1}^{\lfloor \frac{1}{2}\varepsilon \ln n\rfloor}\frac{1}{k}\mathbb{E}[Z^i(n,k)] \\ &\leq \frac{\ln n}{n^2} z^{-\frac{1}{2}\varepsilon \ln n}\sum_{i=1}^{n+1}\mathbb{E}[G_n^i(z)],\end{aligned}$$

for $z \in (0,1)$, where $G_n^i(z) = \sum_{k=0}^{\infty} z^k Z^i(n,k+1)$. We claim the following.

Lemma 3.2. *There exists $z_0 \in (0,1)$ such that we have that*

$$\mathbb{E}[G_n^i(z_0)] \leq e^{\frac{1+z_0}{2}} n^{\frac{1+z_0}{2}}, \quad \textit{for } i \geq 1 \textit{ and } n \geq 1.$$

The proof of the above lemma relies in the recursive structure of the scale-free random tree and for now it is convenient to postpone its proof to Section 3.6. We then consider z_0 such that the result of Lemma 3.2 holds and $0 < \varepsilon < (z_0 - 1)(\ln z_0)^{-1}$. Then

$$\mathbb{E}\left[\frac{\ln n}{2d_n(u,v)}\mathbf{1}_{\{u\neq v\}}\mathbf{1}_{\{d_n(u,v)\leq \frac{1}{2}\varepsilon \ln n\}}\right] \leq e^{\frac{1+z_0}{2}} z_0^{-\frac{1}{2}\varepsilon \ln n} n^{-\frac{1-z_0}{2}} \ln n$$

and therefore, the right-hand side in (3.11) tends to 0 as $n \to \infty$.

Similarly, one can easily check that the uniform random recursive trees fulfill the hypotheses of Corollary 3.2 with $\ell(n) = \ln n$; see Chapter 6 in [8].

3. Merging of regular trees. Our next example provides a method to build trees that fulfill the conditions of Theorem 3.1 and where the random variable ζ_1 in hypothesis (H) is not a constant. Basically, the procedure consists on gluing trees which satisfy the assumptions of Corollary 3.2. In this example, we consider a mixture of regular trees but one may consider other families of trees as well. For a fixed integer $r \geq 1$, let $(d_i)_{i=1}^r$ denote a positive sequence of integers. Next, for $i = 1, \ldots, r$, let $h_i(m) : \mathbb{R}_+ \to \mathbb{R}_+$ be a function with $\lim_{m\to\infty} h_i(m) = \infty$. Moreover, we assume that

$$d_1^{h_1(m)} \sim d_2^{h_2(m)} \sim \cdots \sim d_r^{h_r(m)},$$

when $m \to \infty$. Then, let $T_{n_i}^{(d_i)}$ be a complete d_i-regular tree with height $\lfloor h_i(m)\rfloor$. Since there are d_i^j vertices at distance $j = 0, 1, \ldots, \lfloor h_i(m)\rfloor$ from the root, its size is given by

$$n_i = n_i(m) = d_i(d_i^{\lfloor h_i(m)\rfloor} - 1)/(d_i - 1).$$

In particular, one can check that the assumptions in Theorem 3.1 are fulfilled with $\ell(n_i) = \ln n_i$. We now imagine that we merge all the r regular trees into one common root which leads us to a new tree

$T_n^{(d)}$ of size $n = \sum_{i=1}^r n_i + 1 - r$. Then, we observe that the probability that a vertex of $T_n^{(d)}$ chosen uniformly at random belongs to the tree $T_{n_i}^{(d_i)}$ converges when $m \to \infty$ to $1/r$. Then, one readily checks that this new tree satisfies the hypothesis (H) with $\ell(n) = \ln n$ and ζ_i a random variable uniformly distributed in the set $\{1/\ln d_1, \dots, 1/\ln d_r\}$. Furthermore, since the number of descendants of each vertex is bounded, it is not difficult to see that also fulfills the condition (H'). Therefore, Theorem 3.1 implies that $n^{-1} \ln n \, \mathrm{Cut}(T_n^{(d)})$ converges in distribution in the sense of pointed Gromov-Prokhorov to the element $I_{\mu^{(d)}}$ of $\mathbb{M}$, which corresponds to the interval $[0, a)$, pointed at 0, equipped with the Euclidean distance, and the probability measure $\mu^{(d)}$ given by (3.1) with $\lambda(t) = \frac{1}{r}\sum_{i=1}^r e^{-\frac{t}{\ln d_i}}$ for $t \geq 0$.

3.5 Applications

We now present a consequence of Theorem 3.1 which generalizes a result of Kuba and Panholzer [53], and its recent multi-dimensional extension shown by Baur and Bertoin [54] on the isolation of multiple vertices in uniform random recursive trees. Let $u_1, u_2, \dots$ denote a sequence of i.i.d. uniform random variables in $[n] = \{1, \dots, n\}$. We write $Z_{n,j}$ for the number of cuts which are needed to isolate $u_1, \dots, u_j$ in T_n. We have the following convergence which extends Corollary 4 in [54].

Corollary 3.3. *Suppose that (H) and (H') hold with ℓ such that $\ell(n) = o(\sqrt{n})$. Then as $n \to \infty$, we have that*

$$\left(\frac{\ell(n)}{n} Z_{n,j} : j \geq 1\right) \Rightarrow (\max(U_1, U_2, \dots, U_j) : j \geq 1)$$

in the sense of finite-dimensional distributions, where $U_1, U_2, \dots$ is a sequence of i.i.d. random variables with law μ given in (3.1).

Proof. For a fixed integer $j \geq 1$, $u_1, \dots, u_j$ are j independent uniform vertices of T_n, or equivalently, the singletons $\{u_1\}, \dots, \{u_j\}$ form a sequence of j i.i.d. leaves of $\mathrm{Cut}(T_n)$ distributed according the uniform law. Denote by $\mathcal{R}_{n,j}$ the subtree of $\mathrm{Cut}(T_n)$ spanned by its root and j i.i.d. leaves chosen according to the uniform distribution on $[n]$. Similarly, write $\mathcal{R}_j$ for the subtree of I_μ spanned by 0 and j i.i.d. random variables with law μ, say $U_1, \dots, U_j$. We adopt the framework of Aldous [19], and see both reduced trees as combinatorial trees structure with edge lengths. Therefore, Theorem 3.1 entails that $n^{-1}\ell(n)\mathcal{R}_{n,j}$ converges weakly in the sense of Gromov-Prokhorov to $\mathcal{R}_j$ as $n \to \infty$. In particular, we have the convergence of the lengths of those reduced trees,

$$\left(\frac{\ell(n)}{n}|\mathcal{R}_{n,1}|, \dots, \frac{\ell(n)}{n}|\mathcal{R}_{n,j}|\right) \Rightarrow (|\mathcal{R}_1|, \dots, |\mathcal{R}_j|).$$

It is sufficient to observe that $|\mathcal{R}_j| = \max(U_1, \dots, U_j)$. □

In particular, when the hypotheses (H) and (H') hold with $\zeta_1 \equiv 1$, we observe from Corollary 3.2 that the variables $U_1, U_2, \dots$ have the uniform distribution on $[0, 1]$, and moreover, $\frac{\ell(n)}{n} Z_{n,j}$ converges in distribution to a beta$(j, 1)$ random variable.

As another application, for $j \geq 2$ we consider the algorithm for isolating the vertices $u_1, \dots, u_j$ with a slight modification, we discard the emerging tree components which contain at most one of these j vertices. We stop the algorithm when the j vertices are totally disconnected from each other, i.e. lie in j

different tree components. We write $W_{n,2}$ for the number of steps of this algorithm until for the first time $u_1, \dots, u_j$ do not longer belong to the same tree component, moreover $W_{n,3}$ for the number of steps until the first time, the j vertices are spread out over three distinct tree components, and so on, up to $W_{n,j}$, the number of steps until the j vertices are totally disconnected. We have the following consequence of Corollary 3.1, which extends Corollary 4 in [54].

Corollary 3.4. *Suppose that (H) and (H′) hold with ℓ such that $\ell(n) = o(\sqrt{n})$. Then as $n \to \infty$, we have that*

$$\left(\frac{\ell(n)}{n} W_{n,2}, \dots, \frac{\ell(n)}{n} W_{n,j}\right) \Rightarrow \left(U_{(1,j)}, \dots, U_{(j-1,j)}\right),$$

where $U_{(1,j)} \leq U_{(2,j)} \leq \cdots \leq U_{(j-1,j)}$ denote the first $j-1$ order statistics of an i.i.d. sequence $U_1, \dots, U_j$ of random variables with law μ given in (3.1).

Proof. Recall the notation of Corollary 3.1, and write $Y_i^{(n)}$ for the number of cuts which are needed to disconnect the vertex u_i from the root component. We then observe that if we write $Y_{1,j}^{(n)} \leq Y_{2,j}^{(n)} \leq \cdots \leq Y_{j-1,j}^{(n)}$ for the first order statistics of the sequence of random variables $Y_1^{(n)}, \dots, Y_j^{(n)}$, it follows from Proposition 3.3 that

$$\lim_{n\to\infty} \frac{\ell(n)}{n}(W_{n,i} - Y_{i-1,j}^{(n)}) = 0 \qquad \text{in probability.}$$

Therefore, our claim follows immediately from Corollary 3.1. □

As before, when (H) and (H′) hold with $\zeta_1 \equiv 1$, the variables $U_1, U_2, \dots$ have the uniform distribution on $[0,1]$, and then, $\frac{\ell(n)}{n} W_{n,j}$ converges in distribution to a beta$(1,j)$ random variable, and $\frac{\ell(n)}{n} W_{n,j}$ converges in distribution to a beta$(j-1,2)$ law.

3.6 Proof of Lemma 3.2

The purpose of this final section is to establish Lemma 3.2. The proof relies on the recursive structure of the scale-free random trees, and our guiding line is similar to that in [79] and [80]. We recall that we only consider the case when the parameter α of the scale-free random tree is zero, but that the general case can be treated similarly.

Recall that the construction of the scale-free tree starts at $n = 1$ from the tree $T_1^{(0)}$ on $\{1,2\}$ which has a single edge connecting 1 and 2. Suppose that $T_n^{(0)}$ has been constructed for some $n \geq 2$, then conditionally given $T_n^{(0)}$, the tree $T_{n+1}^{(0)}$ is built by adding an edge between the new vertex $n+2$ and a vertex v_n in $T_n^{(0)}$ chosen at random according to the law

$$\mathbb{P}\left(v_n = i | T_n^{(0)}\right) = \frac{g_n(i)}{2n}, \qquad i \in \{1, \dots, n+1\}.$$

where $g_n(i)$ denotes the degree of the vertex i in $T_n^{(0)}$. Let $Z^i(n,k)$ denote the number of vertices at distance $k \geq 0$ from the vertex i after the n-th step. We are interested in the expectation of the generating function

$$G_n^i(z) = \sum_{k=0}^{\infty} Z^i(n, k+1) z^k, \qquad n \geq 1,$$

for $z \in (0,1)$. In particular, $G_n^1(\cdot)$ is the so-called height profile function; see Katona [79, 80] for several results related to this function. To compute $\mathbb{E}[G_n^i(z)]$ we use the evolution process of the construction of $T_n^{(0)}$ and conditional expectation. Let $\mathcal{F}_n$ denote the σ-field generated by the first n steps in the procedure. The number of vertices at distance k from i increases by one or does not change. Then for $n \geq i-1$,

$$\mathbb{E}[Z^i(n+1,1)|\mathcal{F}_n] = (Z^i(n,1)+1)\frac{Z^i(n,1)}{2n} + Z^i(n,1)\left(1 - \frac{Z^i(n,1)}{2n}\right) = \frac{2n+1}{2n} Z^i(n,1),$$

and for $k > 1$ we have

$$\begin{aligned}
&\mathbb{E}[Z^i(n+1,k)|\mathcal{F}_n] \\
&= (Z^i(n,k)+1)\frac{Z^i(n,k)+Z^i(n,k-1)}{2n} + Z^i(n,k)\left(1 - \frac{Z^i(n,k)+Z^i(n,k-1)}{2n}\right) \\
&= \frac{2n+1}{2n} Z^i(n,k) + \frac{1}{2n} Z^i(n,k-1),
\end{aligned}$$

where $Z^1(0,k) = 0$ and $Z^i(i-2,k) = 0$ for $2 \leq i \leq n+1$. Taking the expectation this leads to the recurrence relation

$$\mathbb{E}[G_{n+1}^i(z)] = \frac{2n+1+z}{2n} \mathbb{E}[G_n^i(z)].$$

Since $G_1^1(z) = G_1^2(z) = 1$, the above recursive formula leads to

$$\mathbb{E}[G_n^1(z)] = \mathbb{E}[G_n^2(z)] = \prod_{j=1}^{n-1} \frac{2j+1+z}{2j}, \tag{3.12}$$

and for $3 \leq i \leq n+1$

$$\mathbb{E}[G_n^i(z)] = \left(\prod_{j=i-1}^{n-1} \frac{2j+1+z}{2j}\right) \mathbb{E}[G_{i-1}^i(z)]. \tag{3.13}$$

with the convention that $\prod_{j=n}^{n-1} \frac{2j+1+z}{2j} = 1$. We point out that $G_n^i(z) = 0$ for $n \leq i-2$. We have the following technical result which will be crucial in the proof of Lemma 3.2.

Lemma 3.3. *For $2 \leq i \leq n$, we have that*

$$\begin{aligned}
&\mathbb{E}[G_n^i(z) Z^i(n,1)] \\
&= \left(\prod_{j=i-1}^{n-1} \frac{2j+2+z}{2j}\right) \mathbb{E}[G_{i-1}^i(z)] + \sum_{k=i-1}^{n-1} \left(\prod_{j=k+1}^{n-1} \frac{2j+2+z}{2j}\right) \frac{1}{2k} \mathbb{E}[Z^i(k,1)],
\end{aligned}$$

and

$$\mathbb{E}[G_n^1(z)Z^1(n,1)] = \prod_{j=1}^{n-1}\frac{2j+2+z}{2j} + \sum_{k=1}^{n-1}\left(\prod_{j=k}^{n-1}\frac{2j+2+z}{2j}\right)\frac{1}{2k}\mathbb{E}[Z^1(k,1)].$$

Proof. We only prove the case when $2 \le i \le n$, the case $i=1$ follows exactly by the same argument. For $n \ge i-1 \ge 1$, we observe that $G_{n+1}^i(z) = G_n^i(z) + K_n^i(z)$ where

$$\mathbb{P}(K_n^i(z) = z^{k-1}|\mathcal{F}_n) = \begin{cases} \frac{Z^i(n,k)+Z^i(n,k-1)}{2n} & k>1 \\ \frac{Z^i(n,1)}{2n} & k=1, \end{cases}$$

and $Z^i(n+1,1) = Z^i(n,1) + B_n^i$ where

$$\mathbb{P}(B_n^i = 1|\mathcal{F}_n) = 1 - \mathbb{P}(B_n^i = 0|\mathcal{F}_n) = \frac{Z^i(n,1)}{2n}.$$

This yields

$$\mathbb{E}(K_n^i(z)|\mathcal{F}_n) = \frac{1+z}{2n}G_n^i(z), \quad \text{and} \quad \mathbb{E}(B_n^i|\mathcal{F}_n) = \mathbb{E}(K_n^i(z)B_n^i|\mathcal{F}_n) = \frac{Z^i(n,1)}{2n}.$$

Then, it follows that

$$\begin{aligned} \mathbb{E}[G_{n+1}^i(z)Z^i(n+1,1)] &= \mathbb{E}[(G_n^i(z)+K_n^i(z))(Z^i(n,1)+B_n^i)|\mathcal{F}_n] \\ &= \frac{2n+2+z}{2n}\mathbb{E}[G_n^i(z)Z^i(n,1)] + \frac{1}{2n}\mathbb{E}[Z^i(n,1)]. \end{aligned}$$

Since $Z^i(i-1,1)=1$, this recursive formula yields to our result. □

Next, we observe that for $1 \le i \le n+1$ the variable $Z^i(n,1)$ is the degree of the vertex i after the n-step, which first moment is given by (see [81])

$$\mathbb{E}[Z^1(n,1)] = \prod_{j=1}^{n-1}\frac{2j+1}{2j}, \quad \text{and} \quad \mathbb{E}[Z^i(n,1)] = \prod_{j=i-1}^{n-1}\frac{2j+1}{2j}, \quad \text{for } 2 \le i \le n+1 \tag{3.14}$$

with the convention that $\prod_{j=n}^{n-1}\frac{2j+1}{2j} = 1$.

We recall some technical results that will be useful later on. We have the following well-known inequality,

$$1 + x \le e^x, \qquad x \in \mathbb{R}. \tag{3.15}$$

Then, we can easily deduce that

$$\prod_{j=i-1}^{n-1}\frac{2j+2+z}{2j} \le e^{\frac{2+z}{2}}\left(\frac{n-1}{i-1}\right)^{\frac{2+z}{2}} \quad \text{and} \quad \prod_{j=i-1}^{n-1}\frac{2j+1}{2j} \le e^{\frac{1}{2}}\left(\frac{n-1}{i-1}\right), \tag{3.16}$$

for $2 \leq i \leq n$. We recall also that by the Euler-Maclaurin formula we have that

$$\sum_{j=1}^{n} \left(\frac{1}{j}\right)^{s} = \left(\frac{1}{n}\right)^{s-1} + s\int_{1}^{n} \frac{\lfloor x \rfloor}{x^{s+1}} \mathrm{d}x, \quad \text{with} \quad s \in \mathbb{R} \setminus \{1\},$$

for $n \geq 1$. Then,

$$\sum_{j=1}^{n} \left(\frac{1}{j}\right)^{s} \leq \left(1 + \frac{s}{1-s}\right) n^{1-s}, \quad \text{for} \;\; s \in (0,1), \tag{3.17}$$

and

$$\sum_{j=1}^{n} \left(\frac{1}{j}\right)^{s} \leq \frac{s}{s-1}, \quad \text{for} \;\; s > 1. \tag{3.18}$$

Lemma 3.4. *There exists $z_0 \in (0,1)$ such that*

$$\mathbb{E}[G^{i}_{i-1}(z_0)] \leq (i-1)^{\frac{1+z_0}{2}}, \quad \textit{for} \;\; i \geq 2. \tag{3.19}$$

Proof. First, we focus on finding the correct z_0. For $i \geq 4$, let v_i be the parent of the vertex i which is distributed according to the law

$$\mathbb{P}\left(v_i = j | T^{(0)}_{i-2}\right) = \frac{Z^{j}(i-2,1)}{2(i-2)}, \quad j \in \{1, 2, \ldots, i-1\}.$$

Then, we have that

$$\begin{aligned} \mathbb{E}[G^{i}_{i-1}(z)] &= 1 + z\mathbb{E}[G^{v_i}_{i-2}(z)] \\ &= 1 + z\sum_{j=1}^{i-1} \mathbb{E}[G^{j}_{i-2}(z)\mathbf{1}_{\{v_i=j\}}] \\ &= 1 + \frac{z}{2(i-2)} \sum_{j=1}^{i-1} \mathbb{E}[G^{j}_{i-2}(z) Z^{j}(i-2,1)]. \end{aligned} \tag{3.20}$$

We observe that Lemma 3.3, (3.14) and (3.16) imply after some computations that

$$\begin{aligned} &\mathbb{E}[G^{j}_{n-2}(z)Z^{j}(n-2,1)] \\ &\leq e^{\frac{2+z}{2}} \left(\frac{n-3}{j-1}\right)^{\frac{2+z}{2}} \left(\mathbb{E}[G^{j}_{j-1}(z)] + \frac{e^{\frac{1}{2}}}{2}(j-1)^{-\frac{3+z}{2}} \sum_{k=j-1}^{n-4} \left(\frac{1}{k}\right)^{\frac{3+z}{2}}\right) + \frac{1}{2}e^{\frac{1}{2}} \left(\frac{1}{(n-3)(j-1)}\right)^{\frac{1}{2}} \end{aligned}$$

for $2 \leq j \leq n-3$. Then the inequalities (3.17) and (3.18) imply that

$$\begin{aligned} &\sum_{j=2}^{n-3} \mathbb{E}[G^{j}_{n-2}(z)Z^{j}(n-2,1)] \\ &\leq e^{\frac{2+z}{2}} (n-3)^{\frac{2+z}{2}} \left(\sum_{j=2}^{n-3} \left(\frac{1}{j-1}\right)^{\frac{2+z}{2}} \mathbb{E}[G^{j}_{j-1}(z)] + e^{\frac{1}{2}} \frac{3+z}{1+z}(n-3)^{\frac{1}{2}}\right) + \frac{e^{\frac{1}{2}}}{(n-3)^{\frac{1}{2}}}, \end{aligned} \tag{3.21}$$

for $n \geq 5$. Similarly, one gets that

$$\mathbb{E}[G^1_{n-2}(z)Z^1(n-2,1)] \leq e^{\frac{2+z}{2}}(n-3)^{\frac{2+z}{2}} + \frac{e^{\frac{3+z}{2}}}{2}\frac{3+z}{1+z}(n-3)^{\frac{2+z}{2}} \tag{3.22}$$

and

$$\mathbb{E}[G^{n-2}_{n-2}(z)Z^{n-2}(n-2,1)] \leq e^{\frac{2+z}{2}}\mathbb{E}[G^{n-2}_{n-3}(z)] + \frac{1}{2}(n-3)^{-1}, \tag{3.23}$$

for $n \geq 4$. Next, we define the functions

$$A^1_n(z) = \left(e^{\frac{2+z}{2}} + \frac{e^{\frac{1+z}{2}}}{2}\frac{3+z}{1+z}\right)(n-3)^{-\frac{1}{2}}, \quad A^2_n(z) = \left(e^{\frac{2+z}{2}} + \frac{1}{2}(n-3)^{-\frac{3+z}{2}}\right)(n-3)^{-1}$$

and

$$A^3_n(z) = 2e^{\frac{2+z}{2}} + e^{\frac{1}{2}}(n-3)^{-\frac{4+z}{2}} + e^{\frac{3+z}{2}}\frac{3+z}{1+z},$$

for $n \geq 4$ and $z \in (0,1)$. Then one can find $z_0 \in (0,1)$ such that

$$3^{-\frac{1+z_0}{2}} + \frac{z_0}{2}\left(A^1_4(z_0) + A^2_4(z_0) + A^3_4(z_0) + \frac{1}{2}\right) \leq 1.$$

Now, we proceed to prove by induction (3.19) with $z_0 \in (0,1)$ such that the previous inequality is satisfied. For $i = 2, 3$, it must be clear since

$$\mathbb{E}[G^2_1(z_0)] = 1 \quad \text{and} \quad \mathbb{E}[G^3_2(z_0)] = 1 + z_0.$$

Suppose that it is true for $i = n-1 \geq 2$. We observe from (3.20) and the inequalities (3.21), (3.22) and (3.23) that

$$\begin{aligned} \mathbb{E}[G^n_{n-1}(z_0)] &\leq 1 + (n-1)^{\frac{1+z_0}{2}}\frac{z_0}{2}\left(A^1_n(z_0) + A^2_n(z_0) + \frac{1}{2} + A^3_n(z_0)\right) \\ &\leq (n-1)^{\frac{1+z_0}{2}}\left(3^{-\frac{1+z_0}{2}} + \frac{z_0}{2}\left(\frac{1}{2} + A^1_4(z_0) + A^2_4(z_0) + A^3_4(z_0)\right)\right) \\ &\leq (n-1)^{\frac{1+z_0}{2}}, \end{aligned}$$

the second inequality is because the functions $A^1_n(\cdot)$, $A^2_n(\cdot)$ and $A^1_n(\cdot)$ are decreasing with respect to n and the last one is by our choice of z_0. □

Finally, we have all the ingredients to prove Lemma 3.2.

Proof of Lemma 3.2. We deduce from the inequality (3.15) that for $n \geq 2$ we have

$$\prod_{j=i-1}^{n-1}\frac{2j+1+z}{2j} \leq e^{\frac{1+z}{2}}\left(\frac{n-1}{i-1}\right)^{\frac{1+z}{2}} \quad \text{for } i \geq 2.$$

We consider $z_0 \in (0, 1)$ such that equation (3.19) in Lemma 3.4 is satisfied. Then from (3.12) and (3.13) we have that

$$\mathbb{E}[G_n^1(z_0)] \leq e^{\frac{1+z_0}{2}} n^{\frac{1+z_0}{2}} \quad \text{for } i \geq 1 \text{ and } n \geq 1,$$

which is our claim. $\square$

CHAPTER 4

Scaling limits for multitype Galton-Watson trees

"I don't want whatever I want. Nobody does. Not really. What kind of fun would it be if I just got everything I ever wanted just like that, and it didn't mean anything? What then?"

— Neil Gaiman, Coraline

In this chapter, we study the behavior of large multitype Galton-Watson trees, as presented in Section 1.4. This work is based on the article [4].

4.1 Introduction

In [19, 61], Aldous introduced the continuum random tree as the limit of rescaled Galton-Watson (GW) trees conditioned on the total progeny, in the case where the offspring distribution has finite variance. More precisely, he proved that the properly rescaled contour process of the conditioned GW tree converges in the functional sense to the normalized Brownian excursion. The latter codes the continuum random tree. This result has motivated the study of the convergence of other rescaled paths obtained from GW trees such as the Lukasiewicz path and the height process. Duquesne and Le Gall [62] showed that the concatenation of rescaled height processes (rescaled contour functions and Lukasiewicz path) converge in distribution to the continuous-time height process associated to a spectrally positive Lévy process. In particular, when the offspring distribution belongs to the domain of attraction of a stable law of index $\alpha \in (1, 2]$, Duquesne [57] showed that the height processes of GW trees conditioned on its total progeny converge in distribution to the normalized excursion of the continuous-time height process associated with a strictly stable spectrally positive Lévy process of index α. Later, Miermont [1] extends the previous results to multitype GW trees. Recall that multitype GW trees are a generalization of usual GW trees that describe the genealogy of a population where individuals are differentiated by types that determine their offspring distribution. Miermont establishes the convergence of the rescaled height process of critical multitype GW trees with finitely many types to the reflected Brownian motion, under the hypotheses that the offspring distribution is irreducible and has finite covariance matrix. Moreover, under an additional exponential moment assumption, he also established that conditionally on the number individuals of a given type, the limit is given by the normalized Brownian excursion. de Raphelis [65] has extended the unconditional result in [1] for multitype GW trees with infinitely many types, under similar assumptions.

In the present chapter, we aim to extend the previous results to critical multitype GW trees with finitely many types whose offspring distribution is still irreducible, but may have infinite variance. More precisely, we are interested in establishing scaling limits for their associated height processes, when the offspring distributions belong to the domain of attraction of a stable law where the stability indices may differ. This will lead us to modify and extend the results of Miermont in [1].

In the rest of the introduction, we will describe our setting more precisely and give the exact definition of multitype GW trees. We then provide the main assumptions on the offspring distribution in Section 4.1.2. This will enable us to state our main results in Section 4.4.2.

4.1.1 Multitype plane trees and forests

We recall the standard formalism for family trees. Let U be the set of all labels:

$$U = \bigcup_{n=0}^{\infty} \mathbb{N}^n,$$

where $\mathbb{N} = \{1, 2, \dots\}$ and with the convention $\mathbb{N}^0 = \{\varnothing\}$. An element of U is a sequence $u = u_1 \cdots u_j$ of positive integers, and we call $|u| = j$ the length of u (with the convention $|\varnothing| = 0$). If $u = u_1 \cdots u_j$ and $v = v_1 \cdots v_k$ belong to U, we write $uv = u_1 \cdots u_j v_1 \cdots v_k$ for the concatenation of u and v. In particular, note that $u\varnothing = \varnothing u = u$. For $u \in U$ and $A \subseteq U$, we let $uA = \{uv : v \in A\}$, and we say that u is a prefix (or ancestor) of v if $v \in uU$, in which case we write $u \vdash v$. Recall that the set U comes with a natural lexicographical order $\prec$, such that $u \prec v$ if and only if either $u \vdash v$, or $u = wu'$, $v = wv'$ with nonempty words u', v' such that $u'_1 < v'_1$.

A rooted planar tree $\mathbf{t}$ is a finite subset of U which satisfies the following conditions:

I. $\varnothing \in \mathbf{t}$, we called it the root of $\mathbf{t}$.

II. For $u \in U$ and $i \in \mathbb{N}$, if $ui \in \mathbf{t}$ then $u \in \mathbf{t}$, and $uj \in \mathbf{t}$ for every $1 \leq j \leq i$.

We let $\mathbb{T}$ be the set of all rooted planar trees. We call vertices (or individuals) the elements of a tree $\mathbf{t} \in \mathbb{T}$, the length $|u|$ is called the height of $u \in \mathbf{t}$. We write $c_{\mathbf{t}}(u) = \max\{i \in \mathbb{Z}_+ : ui \in \mathbf{t}\}$ for the number of children of u. The vertices of $\mathbf{t}$ with no children are called leaves. For $\mathbf{t}$ a planar tree and $u \in \mathbf{t}$, we let $\mathbf{t}_u = \{v \in U : uv \in \mathbf{t}\}$ be the subtree of $\mathbf{t}$ rooted at u, which is itself a tree. The remaining part $[\mathbf{t}]_u = \{u\} \cup (\mathbf{t} \setminus u\mathbf{t}_u)$ is called the subtree of $\mathbf{t}$ pruned at u. The lexicographical order $\prec$ will be called the depth first order on $\mathbf{t}$.

In addition to trees, we are also interested in forests. A forest $\mathbf{f}$ is a nonempty subset of U of the form

$$\mathbf{f} = \bigcup_k k\mathbf{t}_{(k)},$$

where $(\mathbf{t}_{(k)})$ is a finite or infinite sequence of trees, which are called the components of $\mathbf{f}$. In words, a forest may be thought of as a rooted tree where the vertices at height one are the roots of the forest components. We let $\mathbb{F}$ be the set of rooted planar forests. For $\mathbf{f} \in \mathbb{F}$, we define the subtree $\mathbf{f}_u = \{v \in U :$

$uv \in \mathbf{f}\} \in \mathbb{T}$ if $u \in \mathbf{f}$, and $\mathbf{f}_u = \emptyset$ otherwise. Also, let $[\mathbf{f}]_u = \{u\} \cup (\mathbf{f} \setminus u\mathbf{f}_u) \in \mathbb{F}$. With this notation, we observe that the tree components of $\mathbf{f}$ are $\mathbf{f}_1, \mathbf{f}_2, \dots$. We let $c_{\mathbf{f}}(u)$ be the number of children of $u \in \mathbf{f}$. In particular, $c_{\mathbf{f}}(\varnothing) \in \mathbb{N} \cup \{\infty\}$ is the number of components of $\mathbf{f}$. We call $|u| - 1$ the height of $u \in \mathbf{f}$. Notice that that notion of height differs from the convention on trees because we want the roots of the forest components to be at height 0.

Let $d \in \mathbb{N}$, we call $[d] = \{1, \dots, d\}$ the set of types. A d-type planar tree, or simply a multitype tree is a pair $(\mathbf{t}, e_{\mathbf{t}})$, where $\mathbf{t} \in \mathbb{T}$ and $e_{\mathbf{t}} : \mathbf{t} \to [d]$ is a function such that $e_{\mathbf{t}}(u)$ corresponds to the type of a vertex $u \in \mathbf{t}$. We let $\mathbb{T}^{(d)}$ be the set of d-type rooted planar trees. For $i \in [d]$, we write $c^{(i)}_{\mathbf{t}}(u) = \max\{j \in \mathbb{Z}_+ : uj \in \mathbf{t} \text{ and } e_{\mathbf{t}}(uj) = i\}$ for the number of offspring of type i of $u \in \mathbf{t}$. Then, $c_{\mathbf{t}}(u) = \sum_{i \in [d]} c^{(i)}_{\mathbf{t}}(u)$ is the total number of children of $u \in \mathbf{t}$. Analogous definitions hold for d-type rooted planar forests $(\mathbf{f}, e_{\mathbf{f}})$, whose set will be denoted by $\mathbb{F}^{(d)}$. For sake of simplicity, we shall frequently denote the type functions $e_{\mathbf{t}}$, $e_{\mathbf{f}}$ by e when it is free of ambiguity, and will even denote elements of $\mathbb{T}^{(d)}$, $\mathbb{F}^{(d)}$ by $\mathbf{t}$ or $\mathbf{f}$, without mentioning e. Moreover, it will be understood then that $\mathbf{t}_u$, $\mathbf{f}_u$, $[\mathbf{t}]_u$, $[\mathbf{f}]_u$ are marked with the appropriate function.

Finally, for $\mathbf{t} \in \mathbb{T}^{(d)}$ and $i \in [d]$, we let $\mathbf{t}^{(i)} = \{u \in \mathbf{t} : e_{\mathbf{t}}(u) = i\}$ be the set of vertices on $\mathbf{t}$ bearing the type i, and $\mathbf{f}^{(i)}$ the corresponding notation for the forest $\mathbf{f} \in \mathbb{F}^{(d)}$.

4.1.2 Multitype offspring distributions

We set $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ and $d \in \mathbb{N}$. A d-type offspring distribution $\boldsymbol{\mu} = (\mu^{(1)}, \dots, \mu^{(d)})$ is a family of distributions on the space $\mathbb{Z}_+^d$ of integer-valued non-negative sequences of length d. It will be useful to introduce the Laplace transforms $\varphi = (\varphi^{(1)}, \dots, \varphi^{(d)})$ of $\boldsymbol{\mu}$ by

$$\varphi^{(i)}(\mathbf{s}) = \sum_{\mathbf{z} \in \mathbb{Z}_+^d} \mu^{(i)}(\{\mathbf{z}\}) \exp(-\langle \mathbf{z}, \mathbf{s} \rangle), \qquad \text{for } i \in [d],$$

where $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{R}_+^d$ and $\langle x, y \rangle$ is the usual scalar product of two vectors $x, y \in \mathbb{R}^d$. We let $\mathbf{0}$ be the vector of $\mathbb{R}_+^d$ with all components equal to 0. Then, for $i, j \in [d]$, we define the quantity

$$m_{ij} = -\frac{\partial \varphi^{(i)}}{\partial s_j}(\mathbf{0}) = \sum_{\mathbf{z} \in \mathbb{Z}_+^d} z_j \mu^{(i)}(\{\mathbf{z}\})$$

that corresponds to the mean number of children of type j, given by an individual of type i. We let $\mathbf{M} := (m_{ij})_{i,j \in [d]}$ be the mean matrix of $\boldsymbol{\mu}$, and $\mathbf{m}_i = (m_{i1}, \dots, m_{id}) \in \mathbb{R}_+^d$ be the mean vector of the measure $\mu^{(i)}$.

We say that a measure $\boldsymbol{\mu}$ on $\mathbb{Z}_+^d$ is non-degenerate, if there exists at least one $i \in [d]$ so that

$$\mu^{(i)}\left(\left\{\mathbf{z} \in \mathbb{Z}_+^d : \sum_{j=1}^{d} z_j \neq 1\right\}\right) > 0.$$

The offspring distribution that we consider in this work are assumed to be non-degenerate in order to avoid cases which will lead to infinite linear trees.

Definition 4.1. *The mean matrix (or the offspring distribution $\boldsymbol{\mu}$) is called irreducible, if for every $i, j \in [d]$, there is some $n \in \mathbb{N}$ so that $m_{ij}^{(n)} > 0$, where $m_{ij}^{(n)}$ is the ij-entry of the matrix $\mathbf{M}^n$.*

Recall also that if $\mathbf{M}$ is irreducible, then according to the Perron-Frobenius theorem, $\mathbf{M}$ admits a unique eigenvalue ρ which is simple, positive and with maximal modulus. Furthermore, the corresponding right and left eigenvectors can be chosen positive and we call them $\mathbf{a} = (a_1, \ldots, a_d)$ and $\mathbf{b} = (b_1, \ldots, b_d)$ respectively, and normalize them such that $\langle \mathbf{a}, 1\rangle = \langle \mathbf{a}, \mathbf{b}\rangle = 1$; see Chapter V of [17]. We then say that $\boldsymbol{\mu}$ is sub-critical if $\rho < 1$, critical $\rho = 1$ and supercritical if $\rho > 1$.

Main assumptions. Throughout this work, we consider an offspring distribution $\boldsymbol{\mu} = \left(\mu^{(1)}, \ldots, \mu^{(d)}\right)$ on $\mathbb{Z}_+^d$ satisfying the following conditions:

($\mathbf{H}_1$) $\boldsymbol{\mu}$ is irreducible, non-degenerate and critical.

($\mathbf{H}_2$.1) Let Δ be a nonempty subset of $[d]$. For every $i \in \Delta$, there exists $\alpha_i \in (1, 2]$ such that the Laplace transform of $\mu^{(i)}$ satisfies

$$\psi^{(i)}(\mathbf{s}) := -\log \varphi^{(i)}(\mathbf{s}) = \langle \mathbf{m}_i, \mathbf{s}\rangle + |\mathbf{s}|^{\alpha_i}\Theta^{(i)}\left(\mathbf{s}/|\mathbf{s}|\right) + o(|\mathbf{s}|^{\alpha_i}), \quad \text{as } |\mathbf{s}| \downarrow 0,$$

for $\mathbf{s} \in \mathbb{R}_+^d$ and where

$$\Theta^{(i)}(\mathbf{s}) = \int_{\mathbf{S}^d} |\langle \mathbf{s}, \mathbf{y}\rangle|^{\alpha_i} \lambda_i(\mathrm{d}\mathbf{y}),$$

with λ_i a finite Borel non-zero measure on $\mathbf{S}^d = \{\mathbf{y} \in \mathbb{R}^d : |\mathbf{y}| = 1\}$ such that for $\alpha_i \in (1, 2)$, λ_i has support in $\{\mathbf{y} \in \mathbb{R}_+^d : |\mathbf{y}| = 1\}$. We write $|\cdot|$ for the Euclidean norm.

($\mathbf{H}_2$.2) For $i \in [d] \setminus \Delta$, the Laplace transform of $\mu^{(i)}$ satisfies

$$\psi^{(i)}(\mathbf{s}) := -\log \varphi^{(i)}(\mathbf{s}) = \langle \mathbf{m}_i, \mathbf{s}\rangle + o(|\mathbf{s}|^{\alpha_i}), \quad \text{as } |\mathbf{s}| \downarrow 0.$$

where $\alpha_i = \min_{j \in \Delta} \alpha_j$.

Let us comment on these assumptions:

1. We notice that criticality, hypothesis ($\mathbf{H}_1$), implies finiteness of all coefficients of the mean matrix $\mathbf{M}$.

2. For $i \in [d]$, we say that $\mu^{(i)}$ has finite variance when

$$\frac{\partial^2 \varphi^{(i)}}{\partial s_j \partial s_k}(\mathbf{0}) < \infty, \quad \text{for } j, k \in [d].$$

We then write $\mathbf{Q}^{(i)}$ for its covariance matrix. In particular, when $\mu^{(i)}$ satisfies the condition ($\mathbf{H}_2$.1) with $\alpha_i = 2$, one can easily verify that it possess finite variance and that it does not have variance when $\alpha_i \in (1,2)$. This shows that our assumptions on the offspring distribution are less restrictive than the ones made in [1], where the author assumes finiteness of the covariance matrices.

3. In the case when $\mu^{(i)}$ has finite variance, one can consider a measure λ_i on $\mathbf{S}^d$ such that

$$\Theta^{(i)}(\mathbf{s}) = \langle \mathbf{s}, \mathbf{Q}^{(i)}\mathbf{s}\rangle, \qquad \mathbf{s} \in \mathbb{R}_+^d;$$

see for example Section 2.4 of Samorodnitsky and Taqqu [67].

4. Let $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots$ be a sequence of i.i.d. random variables on $\mathbb{Z}_+^d$ with common distribution $\mu^{(i)}$ satisfying $(\mathbf{H}_2)$. We observe that

$$-\log \mathbb{E}\left[\exp\left(-\left\langle \frac{1}{n^{1/\alpha_i}} \sum_{k=1}^{n} (\boldsymbol{\xi}_k - \mathbf{m}_i), \mathbf{s} \right\rangle\right)\right] \underset{n\to\infty}{\to} |\mathbf{s}|^{\alpha_i}\Theta^{(i)}(\mathbf{s}/|\mathbf{s}|), \qquad \mathbf{s} \in \mathbb{R}_+^d, \tag{4.1}$$

Then, we conclude that

$$\frac{1}{n^{1/\alpha_i}} \sum_{k=1}^{n} (\boldsymbol{\xi}_k - \mathbf{m}_i) \xrightarrow[n\to\infty]{d} \mathbf{Y}_{\alpha_i}, \tag{4.2}$$

where $\mathbf{Y}_{\alpha_i}$ is a α_i-stable random vector in $\mathbb{R}_+^d$ whose Laplace exponent satisfies

$$\psi_{\mathbf{Y}_{\alpha_i}}(\mathbf{s}) = |\mathbf{s}|^{\alpha_i}\Theta^{(i)}(\mathbf{s}/|\mathbf{s}|), \qquad \mathbf{s} \in \mathbb{R}_+^d.$$

Sato's book [66] and [67] are good references for background on multivariate stable distributions. On the other hand, we notice from (4.1) that the equation (4.2) is equivalent to the hypothesis $(\mathbf{H}_2.\mathbf{1})$.

5. We point out that in the monotype case, that is $d = 1$, the condition $(\mathbf{H}_2)$ may be thought as the analogous assumption made in [57] and [68], in order to get the convergence of the rescaled monotype GW tree to the continuum stable tree.

6. For $i \in [d] \setminus \Delta$, let $\mu^{(i)}$ be a measure that satisfies the hypothesis $(\mathbf{H}_2.\mathbf{2})$. We can rewrite the expression of its Laplace exponent in the following way

$$\psi^{(i)}(\mathbf{s}) := -\log \varphi^{(i)}(\mathbf{s}) = \langle \mathbf{m}_i, \mathbf{s}\rangle + |\mathbf{s}|^{\alpha_i}\Theta^{(i)}(\mathbf{s}/|\mathbf{s}|) + o(|\mathbf{s}|^{\alpha_i}), \quad \text{as } |\mathbf{s}| \downarrow 0,$$

for $\mathbf{s} \in \mathbb{R}_+^d$ and where

$$\Theta^{(i)}(\mathbf{s}) = \int_{\mathbf{S}^d} |\langle \mathbf{s}, \mathbf{y}\rangle|^{\alpha_i} \lambda_i(d\mathbf{y}),$$

with $\lambda_i \equiv 0$. Recall that $\alpha_i = \min_{j\in\Delta} \alpha_j$ for $i \in [d]\setminus\Delta$. This will be useful for the rest of the chapter.

Finally, let $\underline{\alpha} = \min_{i\in[d]} \alpha_i$ and $\bar{\lambda} = \sum_{i\in[d]} \mathbb{1}_{\{\underline{\alpha}=\alpha_i\}} a_i\lambda_i$. We define

$$\bar{c} = (\langle \mathbf{a}, \boldsymbol{\Theta}(\mathbf{b})\rangle)^{1/\underline{\alpha}} = \left(\int_{\mathbf{S}^d} |\langle \mathbf{b}, \mathbf{y}\rangle|^{\underline{\alpha}} \bar{\lambda}(d\mathbf{y})\right)^{1/\underline{\alpha}},$$

where $\mathbf{\Theta}(\mathbf{s}) = (\Theta^{(1)}(\mathbf{s})\mathbb{1}_{\{\underline{\alpha}=\alpha_1\}}, \dots, \Theta^{(d)}(\mathbf{s})\mathbb{1}_{\{\underline{\alpha}=\alpha_d\}}) \in \mathbb{R}_+^d$, for $\mathbf{s} \in \mathbb{R}_+^d$. We notice that $\bar{c} \not\equiv 0$ due to $(\mathbf{H}_2.\mathbf{1})$. This constant will play a role similar to the constant defined in equation (2) of [1], i.e., it corresponds to the total variance of the offspring distribution $\boldsymbol{\mu}$, when the covariance matrices are finite.

4.1.3 Multitype Galton-Watson trees and forests

Let $\boldsymbol{\mu}$ be a d-type offspring distribution. We define the law $\mathbf{P}_{\boldsymbol{\mu}}^{(i)}$ (or simply $\mathbf{P}^{(i)}$) of a d-type GW tree (or multitype GW tree) rooted at a vertex of type $i \in [d]$ and with offspring distribution $\boldsymbol{\mu}$ by

$$\mathbf{P}^{(i)}(T = \mathbf{t}) = \prod_{u \in \mathbf{t}} \frac{c_{\mathbf{t}}^{(1)}(u)! \dots c_{\mathbf{t}}^{(d)}(u)!}{c_{\mathbf{t}}(u)!} \mu^{(e_{\mathbf{t}}(u))}\left(\left\{c_{\mathbf{t}}^{(d)}(u), \dots, c_{\mathbf{t}}^{(d)}(u)\right\}\right),$$

where $T : \mathbb{T}^{(d)} \to \mathbb{T}^{(d)}$ is the identity map (see e.g., [16], or Miermont [1] for a formal construction of a probability measure on $\mathbb{T}^{(d)}$). In particular, under the criticality assumption, $(\mathbf{H}_1)$, the multitype GW trees with offspring distribution $\boldsymbol{\mu}$ are almost surely finite. Similarly, for $\mathbf{x} = (x_1, \dots, x_r)$ a finite sequence with terms in $[d]$, we define $\mathbf{P}_{\boldsymbol{\mu}}^{\mathbf{x}}$ (or simply $\mathbf{P}^{\mathbf{x}}$) the law of multitype GW forest with roots of type $\mathbf{x}$ and with offspring distribution $\boldsymbol{\mu}$ as the image measure of $\bigotimes_{j=1}^r \mathbf{P}^{(x_j)}$ by the map

$$(\mathbf{t}_{(1)}, \dots, \mathbf{t}_{(r)}) \longmapsto \cup_{k=1}^r k\mathbf{t}_{(k)},$$

i.e., it is the law that makes the identity map $F : \mathbb{F}^{(d)} \to \mathbb{F}^{(d)}$ the random forest whose trees components $F_1, \dots, F_r$ are independent with respective laws $\mathbf{P}^{(x_1)}, \dots, \mathbf{P}^{(x_d)}$. A similar definition holds for an infinite sequence $\mathbf{x} \in [d]^{\mathbb{N}}$.

We then say that a $\mathbb{F}^{(d)}$-value random variable F is a multitype GW forest with offspring distribution $\boldsymbol{\mu}$ and roots of type $\mathbf{x}$ when it has law $\mathbf{P}^{\mathbf{x}}$. Similarly, a $\mathbb{T}^{(d)}$-value random variable T with law $\mathbf{P}^{(i)}$ is a multitype GW tree with offspring distribution $\boldsymbol{\mu}$ and root of type $i \in [d]$.

4.1.4 Main results

In this section, we state our main results on the asymptotic behavior of d-type GW trees with offspring distribution satisfying our main assumptions. In this direction, we first recall the definition of the discrete height process associated to a forest $\mathbf{f} \in \mathbb{F}$.

Let us denote by $\#\mathbf{t}$ the total progeny (or the total number of vertices) of $\mathbf{t}$. Similarly, $\#\mathbf{f}$ represents the total progeny of the forest $\mathbf{f}$. Let $\varnothing = u_{\mathbf{t}}(0) \prec u_{\mathbf{t}}(1) \prec \dots \prec u_{\mathbf{t}}(\#\mathbf{t} - 1)$ be the list of vertices of $\mathbf{t}$ in depth-first order. The height process $H^{\mathbf{t}} = (H_n^{\mathbf{t}}, n \geq 0)$ is defined by $H_n^{\mathbf{t}} = |u_{\mathbf{t}}(n)|$, with the convention that $H_n^{\mathbf{t}} = 0$ for $n \geq \#\mathbf{t}$. For the forest $\mathbf{f}$, we let $1 = u_{\mathbf{f}}(0) \prec u_{\mathbf{f}}(1) \prec \dots \prec u_{\mathbf{f}}(\#\mathbf{f} - 1)$ be the depth-first ordered list of its vertices, and write $H^{\mathbf{f}} = (H_n^{\mathbf{f}}, n \geq 0)$ by $H_n^{\mathbf{f}} = |u_{\mathbf{f}}(n)| - 1$, for $0 \leq n < \#\mathbf{f}$. Detailed description and properties of this object can be found for example in [57].

Let $Y^{(\underline{\alpha})} = (Y_s, s \geq 0)$ be a strictly stable spectrally positive Lévy process with index $\underline{\alpha} \in (1, 2]$ with Laplace exponent

$$\mathbb{E}[\exp(-\lambda Y_s)] = \exp(-s\lambda^{\underline{\alpha}}),$$

for $\lambda \in \mathbb{R}_+$.

We can now state our main result.

Theorem 4.1. *Let F be a d-type GW forest distributed according to $\mathbf{P}^{\mathbf{x}}$, for some arbitrary $\mathbf{x} \in [d]^{\mathbb{N}}$. Then, under $\mathbf{P}^{\mathbf{x}}$, the following convergence in distribution holds for the Skorohod topology on the space $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ of right-continuous functions with left limits:*

$$\left(\frac{1}{n^{1-1/\underline{\alpha}}} H^F_{\lfloor ns \rfloor}, s \geq 0\right) \xrightarrow[n\to\infty]{d} \left(\frac{1}{\bar{c}} H_s, s \geq 0\right),$$

where H stands for the continuous-time height process associated with the strictly stable spectrally positive Lévy process $Y^{(\underline{\alpha})}$.

In particular, we notice that this result implies the convergence in law of the d-type GW forest properly rescaled towards the stable forest of index $\underline{\alpha}$ for the Gromov-Hausdorff topology; see for example Lemma 2.4 of [82]. On the other hand, when $\underline{\alpha} = 2$, it is well-known that $(H_s, s \geq 0)$ is proportional to the reflected Brownian motion. The notion of height process for spectrally positive Lévy process has been studied in great detail in [62].

Next, for $n \geq 0$, we let $\Upsilon^{\mathbf{f}}_n$ be the first letter of $u_{\mathbf{f}}(n)$, with the convention that for $n \geq \#\mathbf{f}$, it equals the number of components of $\mathbf{f}$. In words, $\Upsilon^{\mathbf{f}}_n$ is the index of the tree component to which $u_{\mathbf{f}}(n)$ belongs.

Theorem 4.2. *For $i \in [d]$, let F be a d-type GW forest distributed according to $\mathbf{P}^{\mathbf{i}}$, where $\mathbf{i} = (i, i, \dots)$. Then, under $\mathbf{P}^{\mathbf{i}}$, we have the following convergence in distribution in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$:*

$$\left(\frac{1}{n^{1/\underline{\alpha}}} \Upsilon^F_{\lfloor ns \rfloor}, s \geq 0\right) \xrightarrow[n\to\infty]{d} \left(-\frac{\bar{c}}{b_i} I_s, s \geq 0\right),$$

where I_s is the infimum at time s of the strictly stable spectrally positive Lévy process $Y^{(\underline{\alpha})}$.

Let us explain our approach while we describe the organization for the rest of the chapter. We begin by exposing in Section 4.2.1 the key ingredient, that is, a remarkable decomposition of d-type forests into monotype forests. The plan then is to compare the corresponding height processes of the multitype GW forest and the monotype GW forest, and show that they are close for the Skorohod topology. In this direction, we will need to control the shape of large d-type GW forests. First, we establish in Section 4.2.2 sub-exponential tail bounds for the height and the number of tree components of d-type GW forests that may be of independent interest. Secondly, we estimate in Section 4.2.3 the asymptotic distribution of vertices of the different types. To be a little more precise, Proposition 4.4 provides a convergence of types theorem for multitype GW trees, which extends Theorem 1 (iii) in [1], for the infinite variance case. Roughly speaking, it shows that all types are homogeneously distributed in the limiting tree. We conclude with the proofs of Theorem 4.1 and 4.2 in Section 4.3 by pulling back the known results of Duquesne and Le Gall [62] on the convergence of the rescaled height process of monotype GW forests to the multitype GW forest. Finally, in Section 4.4, we present two applications. The first one is an immediate consequence of Theorem 4.1 and 4.2 which provides information about the maximal height of a vertex in a multitype GW tree. Our second application involves a particular multitype GW tree, known as alternating two-type GW tree which appears frequently in the study of random planar maps.

We establish a conditioned version of Theorem 4.1 for this special tree.

The global structure of the proofs is close to that [1]. Although we will try to make this work as self-contained as possible, we will often refer the reader to this paper when the proofs are readily adaptable, and will rather focus on the new technical ingredients. One difficulty arises from the fact that we are assuming weaker assumptions on the offspring distribution than in [1], we do not assume a finitess of the covariances matrices of the offspring distributions and this forces us to improve some of Miermont's estimates.

4.2 Preliminary results

Through this section unless we specify otherwise, we let F be d-type GW forest with law $\mathbf{P}^{\mathbf{x}}$ where $\mathbf{x} \in [d]^{\mathbb{N}}$ and such that its offspring distribution $\boldsymbol{\mu} = (\mu^{(1)}, \dots, \mu^{(d)})$ satisfies the main assumptions. More precisely, it is important to keep in mind that there is a nonempty subset Δ of $[d]$ such that the family of distributions $(\mu^{(i)})_{i\in\Delta}$ satisfy $(\mathbf{H}_2.\mathbf{1})$ while the remainder $(\mu^{(i)})_{i\in[d]\setminus\Delta}$ fulfills $(\mathbf{H}_2.\mathbf{2})$.

4.2.1 Decomposition of multitype GW forests

In this section, we introduce the projection function $\Pi^{(i)}$ defined by Miermont in [1] that goes from the set of d-types planar forests to the set of monotype planar forests. Roughly speaking, the function $\Pi^{(i)}$ removes all the vertices of type different from i and then it connects the remaining vertices with their most recent common ancestor, preserving the lexicographical order. More precisely, set a d-type forest $\mathbf{f} \in \mathbb{F}^d$ and let $v_1 \prec v_2 \prec \cdots$ be the vertices of $\mathbf{f}^{(i)}$ listed in depth-first order such that all ancestors of v_k have types different from i. They will be the roots of the new forest. We then build a forest $\Pi^{(i)}(\mathbf{f}) = \mathbf{f}'$ with as many tree components as there are elements in $\{v_1, v_2, \dots\}$. Recursively, starting from the set of roots $1, 2, \dots$ of $\mathbf{f}'$, for each $u \in \mathbf{f}'$, we let $v_{u1}, v_{u2}, \dots, v_{uk}$ be vertices of $(v_u \mathbf{f}_{v_u}) \setminus \{v_u\}$ arranged in lexicographical order and such that:

I. They have type i, i.e. $e_{\mathbf{f}}(v_{uj}) = i$ for $1 \le j \le k$,

II. All their ancestors on $(v_u \mathbf{f}_{v_u}) \setminus \{v_u\}$ have types different from i (if any).

Then, we add the vertices $u1, \dots, uk$ to $\mathbf{f}'$ as children of u, and continue iteratively. See Figure 4.1 for an example when $d = 3$.

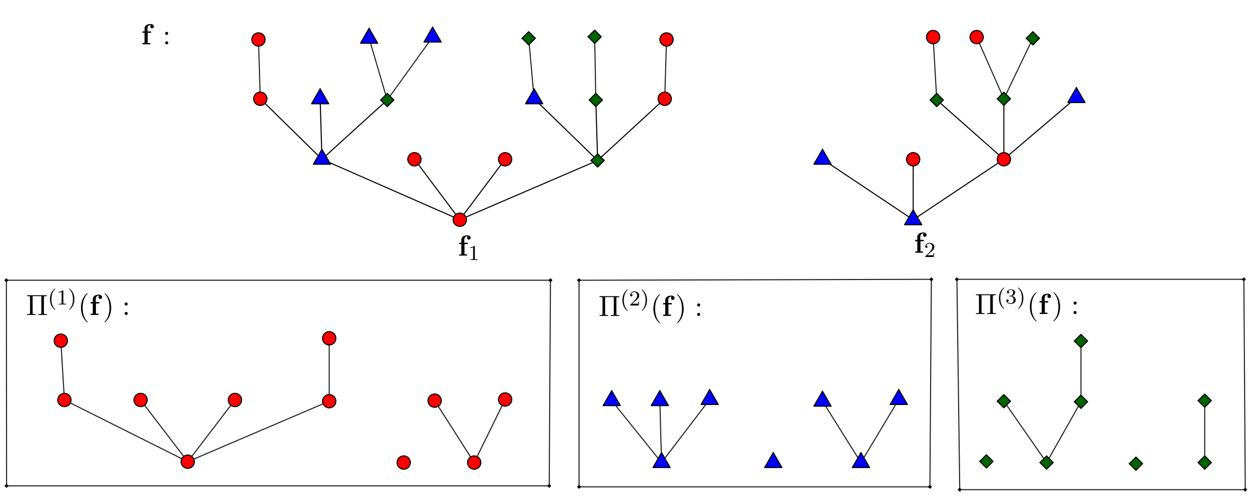

FIGURE 4.1: A realization of the projection $\Pi^{(i)}$ for a three-type planar forest with two tree components, type 1 vertices represented with circles, type 2 vertices with triangles and type 3 vertices with diamonds.

We have the following key result:

Proposition 4.1. *Let $\mathbf{x} \in [d]^{\mathbb{N}}$ and $i \in [d]$. Then, under the law $\mathbf{P}^{\mathbf{x}}$, the forest $\Pi^{(i)}(F)$ is a monotype GW forest with critical non-degenerate offspring distribution $\bar{\mu}^{(i)}$ that is in the domain of attraction of a stable law of index $\underline{\alpha} = \min_{j\in[d]} \alpha_j$. More precisely, the Laplace exponent of $\bar{\mu}^{(i)}$ satisfies*

$$\bar{\psi}^{(i)}(s) = s + \frac{1}{a_i}\left(\frac{\bar{c}}{b_i}s\right)^{\underline{\alpha}} + o(s^{\underline{\alpha}}), \qquad s \downarrow 0,$$

where $s \in \mathbb{R}_+$.

The proof of this proposition is based in an inductive argument that consists in removing types one by one until we are left with a monotype GW forests. More precisely, we suppose that the vertices with type d are removed from the forest $\mathbf{f} \in \mathbb{F}^{(d)}$. We point out that one can delete any other type similarly. We let $v_1 \prec v_2 \prec \dots$ be the vertices of $\mathbf{f}$ listed in depth-first order such that $e_{\mathbf{f}}(v_i) \neq d$ and $e_{\mathbf{f}}(v) = d$ for every $v \vdash v_i$. These are the vertices of $\mathbf{f}$ with type different from d which do not have ancestors of type d. We build a forest $\tilde{\Pi}(\mathbf{f}) = \tilde{\mathbf{f}}$ recursively. We start from the set $\{v_1, v_2, \dots\}$ and for each $v_u \in \tilde{\mathbf{f}}$, we let $v_{u1} \prec \dots \prec v_{uk}$ be the descendants of v_u in $\mathbf{f}$ such that:

I. They have type different from d.

II. For $1 \leq j \leq k$, all the vertices between v_u and v_{uj} have type d (if any).

Then, we add these vertices to $\tilde{\mathbf{f}}$, and continue in an obvious way. We naturally associate the type $e_{\mathbf{f}}$ to the vertices of $\tilde{\Pi}(\mathbf{f})$. In the sequel, we refer to this procedure as the **d- to $(\mathbf{d}-\mathbf{1})$-type operation**.

The following lemma shows that after performing the d-to $(d-1)$-type operation in the multitype GW forest F, we obtain a $(d-1)$-type GW forest which offspring distribution still satisfying our main assumptions. First, we fix some notation. We denote by $\tilde{\mathbf{m}}_d$ the vector in $\mathbb{R}_+^{d-1}$ with entries

$$\tilde{m}_{dk} = \frac{m_{dk}}{1-m_{dd}}, \qquad \text{for } k \in [d-1],$$

and for $j \in [d-1]$, we write $\tilde{\mathbf{m}}_j$ for the vector in $\mathbb{R}_+^{d-1}$ with entries

$$\tilde{m}_{jk} = m_{jk} + \frac{m_{jd}m_{dk}}{1-m_{dd}}, \qquad \text{for } k \in [d-1].$$

We stress that due to the irreducibility assumption on the mean matrix $\mathbf{M}$ of the measure $\boldsymbol{\mu}$, we have that $1 - m_{jj} > 0$ for all $j \in [d]$. Thus, all the previous quantities are finite.

Lemma 4.1. *Let $\mathbf{x} \in [d]^{\mathbb{N}}$. Then, under the law $\mathbf{P}^{\mathbf{x}}$, the forest $\tilde{\Pi}(F)$ is a non-degenerate, irreducible, critical $(d-1)$-type GW forest. Moreover, its offspring distribution $\tilde{\boldsymbol{\mu}} = (\tilde{\mu}^{(1)}, \dots, \tilde{\mu}^{(d-1)})$ has Laplace exponents*

$$\tilde{\psi}^{(j)}(\mathbf{s}) = \langle \tilde{\mathbf{m}}_j, \mathbf{s} \rangle + |\mathbf{s}|^{\tilde{\alpha}_j} \tilde{\Theta}^{(j)}(\mathbf{s}/|\mathbf{s}|) + o(|\mathbf{s}|^{\tilde{\alpha}_j}), \qquad |\mathbf{s}| \downarrow 0,$$

for $j \in [d-1]$, $\mathbf{s} \in \mathbb{R}_+^{d-1}$, $\tilde{\alpha}_j = \min(\alpha_j, \alpha_d)$ and

$$\tilde{\Theta}^{(j)}(\mathbf{s}) = \int_{\mathbf{S}^d} |\langle \mathbf{s}, \tilde{\mathbf{y}} + y_d \tilde{\mathbf{m}}_d \rangle|^{\tilde{\alpha}_j} \tilde{\lambda}_j(\mathrm{d}\mathbf{y}),$$

where $\tilde{\lambda}_j = \mathbb{1}_{\{\tilde{\alpha}_j = \alpha_j\}} \lambda_j + \mathbb{1}_{\{\tilde{\alpha}_j = \alpha_d\}} \frac{m_{jd}}{1-m_{dd}} \lambda_d$, $\mathbf{y} = (y_1, \dots, y_d) \in \mathbb{R}^d$ and $\tilde{\mathbf{y}} = (y_1, \dots, y_{d-1}) \in \mathbb{R}^{d-1}$.

It is important to stress that $\tilde{\lambda}_j \equiv 0$ when $j, d \in [d] \setminus \Delta$, and otherwise it is non-zero (recall the last comment after the introduction of the main assumptions in Section 4.1.2).

Proof. The fact that $\tilde{\Pi}(F)$ is a non-degenerate, irreducible, critical $(d-1)$-type GW forest follows from Lemma 3 (i) in [1]. Moreover, we deduce from this same lemma (see specifically equations (8) and (9) in [1]) that the offspring distribution $\tilde{\boldsymbol{\mu}} = (\tilde{\mu}^{(1)}, \dots, \tilde{\mu}^{(d-1)})$ has Laplace exponents

$$\tilde{\psi}^{(j)}(\mathbf{s}) = \psi^{(j)}(\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s})),$$

for $j \in [d-1]$ and $\mathbf{s} \in \mathbb{R}_+^{d-1}$, where $\tilde{\psi}^{(d)}$ is implicitly defined by

$$\tilde{\psi}^{(d)}(\mathbf{s}) = \psi^{(d)}(\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s})).$$

This is obtained by separating the offspring of each individual with types equal and different from d.

In order to understand the behavior of $\tilde{\psi}^{(j)}$ close to zero, we start by analyzing the one of $\tilde{\psi}^{(d)}$. In this direction, we observe from our main assumptions on the offspring distribution $\boldsymbol{\mu}$ that

$$\begin{aligned}
\tilde{\psi}^{(d)}(\mathbf{s}) &= (1-m_{dd})\langle \tilde{\mathbf{m}}_d, \mathbf{s} \rangle + m_{dd}\tilde{\psi}^{(d)}(\mathbf{s}) + |(\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s}))|^{\alpha_d} \Theta^{(d)} \left(\frac{(\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s}))}{|(\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s}))|} \right) + o(|(\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s}))|^{\alpha_d}) \\
&= \langle \tilde{\mathbf{m}}_d, \mathbf{s} \rangle + \frac{1}{1-m_{dd}} |(\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s}))|^{\alpha_d} \Theta^{(d)} \left(\frac{(\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s}))}{|(\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s}))|} \right) + o(|(\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s}))|^{\alpha_d}),
\end{aligned}$$

as $|\mathbf{s}| \downarrow 0$. We also notice that

$$\tilde{\psi}^{(d)}(\mathbf{s}) = \langle \tilde{\mathbf{m}}_d, \mathbf{s} \rangle + o(|\mathbf{s}|), \qquad \text{as } |\mathbf{s}| \downarrow 0. \tag{4.3}$$

On the one hand, from the above estimate, we know that

$$\langle (\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s})), \mathbf{y} \rangle = \langle \mathbf{s}, \tilde{\mathbf{y}} + y_d \tilde{\mathbf{m}}_d \rangle + y_d o(|\mathbf{s}|), \qquad \text{as } |\mathbf{s}| \downarrow 0,$$

Thus,

$$\begin{aligned} |(\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s}))|^{\alpha_d} \Theta^{(d)} \left(\frac{(\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s}))}{|(\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s}))|} \right) &= \int_{\mathbf{S}^d} |\langle (\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s})), \mathbf{y} \rangle|^{\alpha_d} \lambda_d(\mathrm{d}\mathbf{y}) \\ &= \int_{\mathbf{S}^d} |\langle \mathbf{s}, \tilde{\mathbf{y}} + y_d \tilde{\mathbf{m}}_d \rangle|^{\alpha_d} \lambda_d(\mathrm{d}\mathbf{y}) + o(|\mathbf{s}|^{\alpha_d}) \end{aligned}$$

On the other hand, from (4.3), we have that

$$\langle (\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s})), (\mathbf{s}, \tilde{\psi}^{(d)}(\mathbf{s})) \rangle = \langle \mathbf{s}, \mathbf{s} \rangle + \langle \mathbf{s}, \tilde{\mathbf{m}}_d \rangle^2 + o(|\mathbf{s}|^2), \qquad \text{as } |\mathbf{s}| \downarrow 0.$$

Then, the previous estimates yield that

$$\tilde{\psi}^{(d)}(\mathbf{s}) = \langle \tilde{\mathbf{m}}_d, \mathbf{s} \rangle + \frac{1}{1 - m_{dd}} |\mathbf{s}|^{\alpha_d} \tilde{\Theta}^{(d)} (\mathbf{s}/|\mathbf{s}|) + o(|\mathbf{s}|^{\alpha_d}), \qquad |\mathbf{s}| \downarrow 0, \tag{4.4}$$

where

$$\tilde{\Theta}^{(d)}(\mathbf{s}) = \int_{\mathbf{S}^d} |\langle \mathbf{s}, \tilde{\mathbf{y}} + y_d \tilde{\mathbf{m}}_d \rangle|^{\alpha_d} \lambda_d(\mathrm{d}\mathbf{y}), \qquad \text{for } \mathbf{s} \in \mathbb{R}_+^{d-1}.$$

Finally, from (4.3), (4.4) and our assumption on the Laplace exponent $\psi^{(j)}$, the claim follows by similar computations. $\square$

We notice that after performing the d- to $(d-1)$-type operation, we are left with a non-degenerate, irreducible, critical $(d-1)$-type GW forest whose offspring distribution $\tilde{\boldsymbol{\mu}}$ has mean matrix $\tilde{\mathbf{M}} = (\tilde{m}_{jk})_{j,k \in [d-1]}$. Lemma 4.1 shows that this matrix has spectral radius 1 and moreover, it is not difficult to check that its left and right 1-eigenvectors $\tilde{\mathbf{a}}$, $\tilde{\mathbf{b}}$ satisfying $\langle \tilde{\mathbf{a}}, \mathbf{1} \rangle = \langle \tilde{\mathbf{a}}, \tilde{\mathbf{b}} \rangle = 1$ are given by

$$\tilde{\mathbf{a}} = \frac{1}{1 - a_d}(a_1, \dots, a_d) \quad \text{and} \quad \tilde{\mathbf{b}} = \frac{1 - a_d}{1 - a_d b_d}(b_1, \dots, b_d).$$

We are now able to establish Proposition 4.1.

Proof of Proposition 4.1. The fact that $\Pi^{(i)}(F)$ is a monotype GW forest with critical non-degenerate offspring distribution is a consequence of Lemma 4.1 by following exactly the same argument as the proof of Proposition 4 (i) in [1]. Roughly speaking, the idea is to remove the types different from i one by one through the d- to $(d-1)$-type operation, and noticing that the hypotheses of the GW forest under consideration are conserved at every step until we are left with a critical non-degenerate monotype GW forest. Thus, it only remains to show that the offspring distribution of $\Pi^{(i)}(F)$ is in the domain of attraction of a stable law of index $\underline{\alpha} = \min_{j \in [d]} \alpha_j$. We prove this by induction on d, in the case $i = 1$, without losing generality. The case $d=1$ is obvious. So suppose $d \geq 2$. We apply the d- to $(d-1)$-type operation through Lemma 4.1, and use the induction hipothesis to conclude that the offspring distribution $\bar{\mu}^{(1)}$ of $\Pi^{(1)}(F)$ satisfies

$$\bar{\psi}^{(1)}(s) = s + \frac{1}{\tilde{a}_1}\left(\frac{\tilde{c}}{\tilde{b}_1}\right)^{\underline{\alpha}} + o(s^{\underline{\alpha}}), \qquad s \downarrow 0,$$

for $s \in \mathbb{R}_+$ and where $\tilde{c} = (\langle \tilde{\mathbf{a}}, \tilde{\mathbf{\Theta}}(\tilde{\mathbf{b}})\rangle)^{1/\underline{\alpha}}$, with

$$\tilde{\mathbf{\Theta}}(\mathbf{s}) = \left(\tilde{\Theta}^{(1)}(\mathbf{s})\mathbb{1}_{\{\underline{\alpha}=\tilde{\alpha}_1\}}, \dots \tilde{\Theta}^{(d-1)}(\mathbf{s})\mathbb{1}_{\{\underline{\alpha}=\tilde{\alpha}_{d-1}\}}\right),$$

as in Lemma 4.1. On the other habd, we first observe that for $j \in [d-1]$, we have

$$\begin{aligned}
\tilde{\Theta}^{(j)}(\tilde{\mathbf{b}}) &= \int_{\mathbf{S}^d} |\langle \tilde{\mathbf{b}}, \tilde{\mathbf{y}} + y_d \tilde{\mathbf{m}}_d\rangle|^{\tilde{\alpha}_j} \tilde{\lambda}_j(\mathrm{d}\mathbf{y}) \\
&= \int_{\mathbf{S}^d} \left|\langle \tilde{\mathbf{b}}, \tilde{\mathbf{y}}\rangle + y_d \langle \tilde{\mathbf{b}}, \tilde{\mathbf{m}}_d\rangle\right|^{\tilde{\alpha}_j} \tilde{\lambda}_j(\mathrm{d}\mathbf{y}) \\
&= \left(\frac{1-a_d}{1-a_d b_d}\right)^{\tilde{\alpha}_j} \int_{\mathbf{S}^d} \left|\sum_{k=1}^{d-1} b_k y_k + y_d \sum_{k=1}^{d-1} b_k \frac{m_{dk}}{1-m_{dd}}\right|^{\tilde{\alpha}_j} \tilde{\lambda}_j(\mathrm{d}\mathbf{y}) \\
&= \left(\frac{1-a_d}{1-a_d b_d}\right)^{\tilde{\alpha}_j} \tilde{\Theta}^{(j)}(\mathbf{b}),
\end{aligned}$$

where for the last equality, we use the fact the $\mathbf{b}$ is the right 1-eigenvector of the mean matrix $\mathbf{M}$, that is, $\sum_{k\in[d]} b_k m_{dk} = b_d$. Then, from the previous identity, we have that

$$\begin{aligned}
\langle \tilde{\mathbf{a}}, \tilde{\mathbf{\Theta}}(\tilde{\mathbf{b}})\rangle &= \left(\frac{1-a_d}{1-a_d b_d}\right)^{\underline{\alpha}} \left(\sum_{k=1}^{d-1} \tilde{a}_k \Theta^{(k)}(\mathbf{b})\mathbb{1}_{\{\underline{\alpha}=\alpha_k\}} + \Theta^{(d)}(\mathbf{b})\mathbb{1}_{\{\underline{\alpha}=\alpha_d\}} \sum_{k=1}^{d-1} \tilde{a}_k \frac{m_{kd}}{1-m_{dd}}\right) \\
&= \frac{(1-a_d)^{\underline{\alpha}-1}}{(1-a_d b_d)^{\underline{\alpha}}} \langle \mathbf{a}, \mathbf{\Theta}(\mathbf{b})\rangle,
\end{aligned}$$

where in the last equality, we now use that $\mathbf{a}$ is the left 1-eigenvector of the mean matrix $\mathbf{M}$, i.e., $\sum_{k\in[d]} a_k m_{kd} = a_d$. Therefore, the fact that $\bar{\mu}^{(1)}$ belongs to the domain of attraction of a stable law of index $\underline{\alpha}$ follows by induction and the above equality. □

Following Miermont [1], we are interested in keeping the information of the number vertices that we delete during the projection $\Pi^{(i)}$. More precisely, for $\mathbf{f} \in \mathbb{F}^{(d)}$, recall that $\Pi^{(i)}(\mathbf{f})$ is the monotype forest obtained by removing all the vertices with type different from i. Then, for a vertex $u \in \Pi^{(i)}(\mathbf{f})$ with children $u1, \dots, uk$, we let $\mathbf{f}_{v_u}, \mathbf{f}_{v_{u1}}, \dots, \mathbf{f}_{v_{uk}}$ be the subtrees of the original forest $\mathbf{f}$ rooted at $u, u1, \dots, uk$, respectively. Then, we let

$$N_{ij}(u) = \#\left\{w \in \mathbf{f}_{v_u} \setminus \left(\bigcup_{r=1}^{k} \mathbf{f}_{v_{ur}}\right) : e_{\mathbf{f}}(w) = j\right\}, \qquad \text{for } j \in [d]\setminus\{i\},$$

be the number of type j vertices that have been deleted between u and its children. We also let

$$\hat{N}_{ij}(n) = \#\left\{v \in \mathbf{f}_n : e_{\mathbf{f}}(v) = j \ \text{ and } \ e_{\mathbf{f}}(w) \neq i \text{ for all } w \vdash v\right\}, \qquad \text{for } j \in [d]\setminus\{i\},$$

be the number of type j vertices of the n-th tree component of $\mathbf{f}$ that lie below the first layer of type i vertices, i.e. the number of type j vertices of $\mathbf{f}_n$ that do not have ancestors of type i.

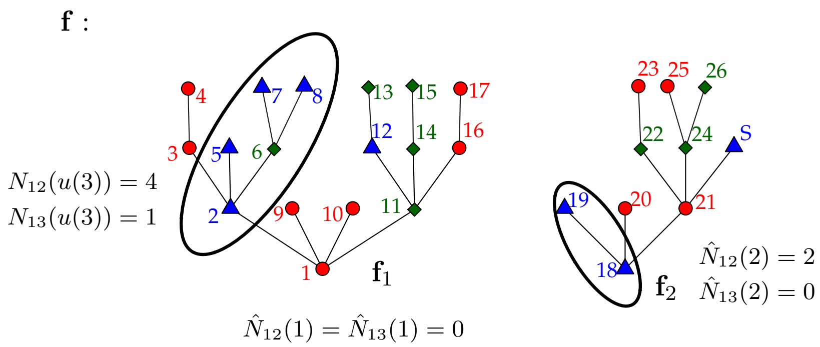

FIGURE 4.2: A representation of the quantities N_{1j} and $\hat{N}_{2j}$, for a three-type planar forest with two tree components, type 1 vertices represented with circles, type 2 vertices with triangles and type 3 vertices with diamonds.

The following proposition provides information about the distribution of the previous quantities.

Proposition 4.2. *Let $1 = u(0) \prec u(1) \prec \cdots \prec u(\#\Pi^{(i)}(\mathbf{f}) - 1)$ be the list of vertices of $\Pi^{(i)}(\mathbf{f})$ in depth-first order and let $\mathbf{x} \in [d]^{\mathbb{N}}$. Then, under the law $\mathbf{P}^{\mathbf{x}}$ and for each $i \in [d]$:*

(i) For every $j \in [d] \backslash \{i\}$, the random variables $(N_{ij}(u(n)), n \geq 0)$ are i.i.d. Moreover, their Laplace exponents satisfy

$$\phi_{ij}(s) := -\log \mathbf{E}^{\mathbf{x}} \left[\exp\left(-sN_{ij}(u(0))\right)\right] = \frac{a_j}{a_i} s + c_{ij} s^{\underline{\alpha}} + o(s^{\underline{\alpha}}), \qquad \text{as } s \downarrow 0,$$

where $s \in \mathbb{R}_+$, $\underline{\alpha} = \min_{j\in[d]} \alpha_j$ and $c_{ij} > 0$ a constant. In particular, $\mathbf{E}^{\mathbf{x}}[N_{ij}(u(0))] = a_j/a_i$.

(ii) For every $j \in [d] \setminus \{i\}$, the random variables $(\hat{N}_{ij}(n), n \geq 1)$ are independent, and their Laplace exponents satisfy

$$\hat{\phi}_{ij}(s) := -\log \mathbf{E}^{\mathbf{x}} \left[\exp\left(-s\hat{N}_{ij}(n)\right)\right] = \left(\hat{c}_{ij} s + \hat{c}'_{ij} s^{\hat{\alpha}_i} + o(s^{\hat{\alpha}_i})\right) \mathbb{1}_{\{x_n \neq i\}}, \qquad \text{as } s \downarrow 0,$$

for $s \in \mathbb{R}_+$, some constants $\hat{c}_{ij} > 0$ and $\hat{c}'_{ij} \geq 0$ (that depends on x_n) and where $\hat{\alpha}_i = \min_{j\in[d]\backslash\{i\}} \alpha_j$.

The idea of the proof is based in a similar induction argument as in the one of Proposition 4.1, by making use of the d- to $(d-1)$-type operation $\tilde{\Pi}$. In this direction, we notice that the left and right 1-eigenvectors $\mathbf{a}$, $\mathbf{b}$ of $\mathbf{M}$ satisfy, for $1 \leq j \leq d$,

$$a_j = \sum_{i=1}^{d} a_i m_{ij} \qquad \text{and} \qquad b_i = \sum_{j=1}^{d} b_j m_{ij}$$

for $1 \leq i \leq d$. In particular, when $d = 2$, a simple computation shows that

$$\frac{a_2}{a_1} = \frac{m_{12}}{1 - m_{22}} = \frac{1 - m_{11}}{m_{21}} \qquad \text{and} \qquad \frac{b_2}{b_1} = \frac{m_{21}}{1 - m_{22}} = \frac{1 - m_{11}}{m_{12}}.$$

This will be useful in a moment.

Proof of Proposition 4.2.

(i) The fact that for every $j \in [d] \setminus \{i\}$, the random variables $(N_{ij}(u(n)), n \geq 0)$ are i.i.d. has been proven in Proposition 4 (ii) of [1]. Basically, this follows from Jagers' theorem on stopping lines [83]. We then focus on the second part of the statement, and for simplicity, we prove this in the case $i = 1$, without losing generality. In this direction, for $\mathbf{f} \in \mathbb{F}^{(d)}$ and $u \in \tilde{\Pi}(\mathbf{f})$, we let $\tilde{N}(u)$ be the number of d-type vertices that have been deleted between u and its children during this procedure. For $j \in [d-1]$, we let $u^{(j)}(0) \prec u^{(j)}(1) \prec \ldots$ be the type j vertices of F arranged in depth-first order. Then, Lemma 3 (ii) in [1] ensures that under $\mathbf{P}^{\mathbf{x}}$, the $d-1$ sequences $(\tilde{N}(u^{(j)}(n)), n \geq 0)$ are independent and formed of i.i.d. elements. Further, their Laplace exponents $\tilde{\phi}^{(j)}$ respectively satisfy

$$\tilde{\phi}^{(j)}(s) = \psi^{(j)}(\mathbf{0}, \tilde{\phi}^{(d)}(s))$$

for $s \in \mathbb{R}_+$, $\mathbf{0}$ the vector of $\mathbb{R}_+^{d-1}$ with all components equal to 0, and where $\tilde{\phi}^{(d)}$ is implicitly given by

$$\tilde{\phi}^{(d)}(s) = s + \psi^{(d)}(\mathbf{0}, \tilde{\phi}^{(d)}(s)). \tag{4.5}$$

Thus, from our main assumptions on the offspring distribution, it is not difficult to check by following the same reasoning as the proof of Lemma 4.1 that

$$\tilde{\phi}^{(j)}(s) = \frac{m_{jd}}{1 - m_{dd}} s + \tilde{c}_{jd} s^{\tilde{\alpha}_j} + o(s^{\tilde{\alpha}_j}), \qquad \text{as } s \downarrow 0,$$

where $\tilde{\alpha}_j = \min(\alpha_j, \alpha_d)$ and the constant $\tilde{c}_{jd} = 0$ if $j, d \in [d] \setminus \Delta$ and $\tilde{c}_{jd} > 0$ otherwise (recall the main assumptions $(\mathbf{H}_2.\mathbf{1})$ and $(\mathbf{H}_2.\mathbf{2})$).

Let now proceed to prove our statement. In the monotype case, $d = 1$, there is nothing to show. For the case $d = 2$, one checks from the previous discussion that the Laplace exponent of $N_{12}(u(0))$ satisfies

$$\phi_{12}(s) = \frac{m_{12}}{1 - m_{22}} s + \tilde{c}_{12} s^{\tilde{\alpha}_1} + o(s^{\tilde{\alpha}_1}), \qquad \text{as } s \downarrow 0.$$

On the other hand, we know that that $m_{12}/(1 - m_{22}) = a_2/a_1$.

We now consider case $d \geq 3$. We apply the operation $\tilde{\Pi}$, $d-2$ times, removing the types $d, d-1, \ldots, 3$ one after the other. We then obtain a two-type GW forest and we observe that the number of type 2 vertices that have only the root as type 1 ancestor is precisely the number of type 2 individuals that are trapped between two generations of $\Pi^{(1)}(F)$. Therefore, in view of the $d = 2$ case above, it is not difficult to see that the Laplace exponent of $N_{12}(u(0))$ satisfies

$$\phi_{12}(s) = \frac{a_2}{a_1} s + c_{12} s^{\underline{\alpha}} + o(s^{\underline{\alpha}}), \qquad \text{as } s \downarrow 0,$$

for some constant $c_{12} > 0$. Finally, our claim follows by symmetry.

(ii) This is obtained by a similar induction argument. We only need to notice that for $i \in [d]$ and $j \in [d] \setminus \{i\}$, $\hat{N}_{ij}(n) = 0$ when $x_n = i$.

□

4.2.2 Sub-exponential Bounds

The following lemma gives an exponential control on the height and number of components related to the n first vertices in d-type GW forests. This extends Lemma 4 in [1] which considers the finite variance case. Recall that for a forest $\mathbf{f} \in \mathbb{F}$, we let $1 \prec u_{\mathbf{f}}(0) \prec u_{\mathbf{f}}(1) \prec \cdots \prec u_{\mathbf{f}}(\#\mathbf{f}-1)$ be the depth-first ordered list of its vertices. Recall also that $\Upsilon_n^{\mathbf{f}}$ is the index of the tree component to which $u_{\mathbf{f}}(n)$ belongs.

Lemma 4.2. *There exist two constants $0 < C_1, C_2 < \infty$ (depending only on $\boldsymbol{\mu}$) such that for every $n \in \mathbb{N}$, $\mathbf{x} \in [d]^{\mathbb{N}}$ and $\eta > 0$,*

$$\mathbf{P}^{\mathbf{x}} \left(\max_{0 \leq k \leq n} |u_F(k)| \geq n^{1-1/\underline{\alpha}+\eta} \right) \leq C_1 (n+1) \exp\left(-C_2 n^{\eta}\right)$$

and

$$\mathbf{P}^{\mathbf{x}} \left(\Upsilon_n^F \geq n^{1/\underline{\alpha}+\eta} \right) \leq C_1 \exp\left(-C_2 n^{\eta}\right).$$

Proof. We observe that under $\mathbf{P}^{\mathbf{x}}$ and independently of $\mathbf{x}$, we have that

$$\max_{0 \leq k \leq n} |u_F(k)| \leq \sum_{i \in [d]} \max_{0 \leq k \leq n} \left| u_{\Pi^{(i)}(F)}(k) \right| \quad \text{and} \quad \Upsilon_n^F \leq \sum_{i \in [d]} \Upsilon_n^{\Pi^{(i)}(F)},$$

where each of the forests $\Pi^{(i)}(F)$, for $i \in [d]$, are critical non-degenerate monotype GW forests with offspring distribution in the domain of attraction of a stable law of index $\underline{\alpha} \in (1, 2]$ by Proposition 4.1. Therefore, from the above inequalities, it is enough to prove the result only for the case $d = 1$.

In this direction, let μ be a critical non-degenerate offspring distribution on $\mathbb{Z}_+$, with Laplace exponent given by

$$\psi(s) = s + cs^{\alpha} + o(s^{\alpha}), \qquad \text{as } s \downarrow 0,$$

for $\alpha \in (1, 2]$, $s \in \mathbb{R}_+$ and $c > 0$ a constant. Let $\mathbf{P}$ be the law of a monotype GW forest with an infinite number of components and offspring distribution μ. We then let F be a monotype GW forest with law $\mathbf{P}$.

It is well-known ([62], Section 2.2) that $|u_F(k)| - 1$ has the same distribution as the number of weak records for a random walk with step distribution $\mu(\{\cdot + 1\})$ on $\{-1\} \cup \mathbb{Z}_+$, from time 1 up to time k. We denote by $W = (W_n, n \geq 0)$ such random walk and we also consider that is defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. It is important to point out that W is the well-known Lukasiewicz path associated with the GW forest with offspring distribution μ. By assumption, the step distribution of this random walk is centered and in the domain of attraction of stable law of index $\alpha \in (1, 2]$. That is, $W_n/n^{1/\alpha}$ converges in distribution towards a stable law of index α as $n \to \infty$. We fix $\tau_0 = 0$ and write τ_j, $j \geq 0$, for the time of the j-th weak record of $(W_n, n \geq 0)$. Therefore, from [84] and Theorems 1 and 2 in [85], the sequence of random variables $(\tau_j - \tau_{j-1}, j \geq 1)$ is i.i.d. with Laplace exponent given by

$$\tilde{\kappa}(\lambda) = -\log \mathbb{E}\left[\exp\left(-\lambda \tau_1\right)\right] = \tilde{C}_1 \lambda^{1-1/\alpha} + o(\lambda^{1-1/\alpha}), \qquad \text{as } \lambda \downarrow 0, \tag{4.6}$$

for some constant $\tilde{C}_1 > 0$. We then bound the first probability by

$$\mathbf{P}\left(\max_{0\leq k\leq n}|u_F(k)| \geq n^{1-1/\alpha+\eta}\right) \leq (n+1)\max_{0\leq k\leq n}\mathbf{P}\left(|u_F(k)| \geq n^{1-1/\alpha+\eta}\right).$$

Then, we notice that for $0 \leq k \leq n$ and $m \in \mathbb{N}$, we have that

$$\mathbf{P}\left(|u_F(k)| - 1 \geq m\right) = \mathbb{P}\left(\sum_{j=1}^{m}(\tau_j - \tau_{j-1}) \leq k\right) \leq e\mathbb{E}\left[\exp\left(-\sum_{j=1}^{m}\frac{\tau_j - \tau_{j-1}}{k}\right)\right] \leq \exp\left(1 - m\tilde{\kappa}(1/n)\right),$$

where for the last inequality, we use the monotonicity of $\tilde{\kappa}$. Taking $m = \lceil n^{1-1/\alpha+\eta}\rceil - 1$ and using (4.6), we get the first bound for large n and thus for every n up to tuning the constants C_1, C_2.

The proof for second bound is very similar. For $j \geq 1$, let $\#F_j$ be the number of vertices of the j-th tree component of the forest F. By the Otter-Dwass formula (see, e.g., [63], Chapter 5), under $\mathbf{P}$, $(\#F_i, i \geq 1)$ is a sequence of i.i.d. random variables with common distribution

$$\mathbf{P}\left(\#F_1 = n\right) = n^{-1}\mathbb{P}\left(W_n = -1\right).$$

Using again the fact that the step distribution of $(W_n, n \geq 0)$ is centered and in the domain of attraction of a stable law of index α, we obtain that

$$\mathbf{P}\left(\#F_1 = n\right) = \tilde{C}_2 n^{-1-1/\alpha} + o(n^{-1-1/\alpha}), \qquad \text{as } n \to \infty,$$

where $\tilde{C}_2 > 0$ is some positive constant; see for example Lemma 1 in [68]. Therefore, an Abelian theorem ([84], Theorem XIII.5.5) entails that the Laplace exponent κ of the distribution of $\#F_1$, under $\mathbf{P}$, satisfies

$$\kappa(\lambda) = \tilde{C}_3\lambda^{1/\alpha} + o(\lambda^{1/\alpha}), \qquad \text{as } \lambda \downarrow 0, \tag{4.7}$$

for some constant $\tilde{C}_3 > 0$. Noticing that $\left\{\Upsilon_n^F(n) \geq m\right\} = \left\{\sum_{i=1}^{m-1}\#F_i \leq n\right\}$, the second bound is then obtained analogously as the first one. Finally, we tune up the constants C_1, C_2 so that they match to both cases. □

4.2.3 Convergence of types

In order to compare the height process of the monotype GW forest $\Pi^{(i)}(F)$, $i \in [d]$, with that of the d-type GW forest F, we must estimate the number of vertices of F that stand between a type i vertex of $\Pi^{(i)}(F)$ and one of its descendants. This is the purpose of the following result. Before that, we need some further notation.

Definition 4.2. *We say that a sequence of positive numbers $(z_n, n \geq 0)$ is exponentially bounded if there are positive constants $c, C > 0$ such that $z_n \leq Ce^{-cn^{\epsilon}}$ for some $\varepsilon > 0$ and large enough n. In order to simplify notations and avoid referring to the changing ε's and the constants c and C, we write $z_n = oe(n)$ in this case.*

For a d-type forest $\mathbf{f} \in \mathbb{F}^{(d)}$ and a vertex $u \in \mathbf{f}$, we let $\mathrm{Anc}^u_{\mathbf{f}}(i)$ be the number of type i ancestors of a vertex u. We provide the following key estimate for the height process which is the analogue of Proposition 5 in [1].

Proposition 4.3. *For every $\gamma > 0$ and $\mathbf{x} \in [d]^{\mathbb{N}}$, we have that*

$$\max_{i\in[d]} \mathbf{P}^{\mathbf{x}} \left(\max_{0\le k\le n} \left| H_k^F - \frac{\mathrm{Anc}_F^{u(k)}(i)}{a_i b_i} \right| > n^{1/2-1/2\underline{\alpha}+\gamma} \right) = \mathrm{oe}(n).$$

The proof of this statement follows exactly from the same argument in [1] and a detailed proof would be cumbersome. So, we leave the details to the interedted reader.

On the other hand, observe that the height process of the monotype GW forest $\Pi^{(i)}(F)$ does not visit the vertices of type different from i, in words, it goes faster than the the height process of the d-type GW forest F. Then, in order to slow down the height process of $\Pi^{(i)}(F)$, we must adjust the time. We conclude this section with the following result which takes care of the number of vertices with type different from i that stands between two consecutive type i vertices in $\Pi^{(i)}(F)$. More precisely, for $\mathbf{f} \in \mathbb{F}^{(d)}$ and $n \ge 0$, we let

$$\Lambda_i^{\mathbf{f}}(n) = \#\left\{0 \le k \le n : e_{\mathbf{f}}(u_{\mathbf{f}}(k)) = i\right\}$$

be the number of type i vertices standing before the $(n+1)$-th vertex in depth-first order. We let $u^{(i)}(0) \prec u^{(i)}(1) \prec \dots$ be the type i vertices of $\mathbf{f}$ arranged in depth-first order, and we also consider the quantity $G_i^{\mathbf{f}}(n) = \#\{u \in \mathbf{f} : u \prec u^{(i)}(n)\}$, with the convention $G_i^{\mathbf{f}}(\#\mathbf{f}^{(i)}) = \#\mathbf{f}$. Similar notation holds if we consider trees instead of forests. Recall that $\mathbf{a} = (a_1, \dots, a_d)$ is the left 1-eigenvector of the mean matrix $\mathbf{M}$.

Proposition 4.4. *For $i \in [d]$ and for any $\mathbf{x} \in [d]^{\mathbb{N}}$, under $\mathbf{P}^{\mathbf{x}}$, we have that*

$$\left(\frac{\Lambda_i^F(\lfloor ns \rfloor)}{n}, s \ge 0 \right) \underset{n\to\infty}{\to} (a_i s, s \ge 0),$$

in probability, for the topology of uniform convergence over compact subsets of $\mathbb{R}_+$.

Proof. We only need to prove that for $i \in [d]$, $\varepsilon > 0$ and for any $\mathbf{x} \in [d]^{\mathbb{N}}$, we have that

$$\mathbf{P}^{\mathbf{x}} \left(\left| G_i^F(n) - a_i^{-1} n \right| > \varepsilon n \right) = 0, \tag{4.8}$$

as $n \to \infty$. This will imply the convergence in probability for every rational number s of $G_i^F(\lfloor ns \rfloor) n^{-1}$ towards $a_i^{-1} s$ as $n \to \infty$. Then, an application of Skorohod's representation theorem and a standard diagonal procedure entail that the above convergence holds for the uniform topology over compact subsets of $\mathbb{R}_+$. Finally, one notices that Λ_i^F is the right-continuous inverse function of G_i^F which leads to our statement.

In this direction, for $\mathbf{f} \in \mathbb{F}^{(d)}$, we recall that $\Pi^{(i)}(\mathbf{f})$ denotes the monotype forest obtained after applying the projection function described in Section 4.2.1. Recall that for $k \ge 0$ and $j \in [d] \setminus \{i\}$, $N_{ij}(k) := N_{ij}(u^{(i)}(k))$ denotes the number of type j vertices that have been deleted between $u(k)$ and its children during the operation $\Pi^{(i)}$. Similarly, we define the quantity $N'_{ij}(k)$ which counts only the type

j vertices that come before $u^{(i)}(n)$ in depth-first order. Since $\sum_{j\neq i} a_j/a_i = 1 - 1/a_i$, we notice that

$$G_i^{\mathbf{f}}(n) - a_i^{-1} n = \sum_{j\neq i} \left(R_1^{\mathbf{f}}(j;n) + R_2^{\mathbf{f}}(j;n) + R_3^{\mathbf{f}}(j;n) \right), \tag{4.9}$$

for $n \geq 0$ and where for $j \in [d] \setminus \{i\}$,

$$R_1^{\mathbf{f}}(j;n) = \sum_{k=0}^{n-1} \left(N'_{ij}(k) - N_{ij}(k) \right) \mathbb{1}_{\{u^{(i)}(k) \vdash u^{(i)}(n)\}}, \qquad R_2^{\mathbf{f}}(j;n) = \sum_{k=1}^{\Upsilon_n^{\mathbf{f}}} \hat{N}_{ij}(k),$$

and

$$R_3^{\mathbf{f}}(j;n) = \sum_{k=0}^{n-1} \left(N_{ij}(k) - a_j/a_i \right).$$

We next estimate the probability that these tree terms are large, when we consider a d-type GW forest. We fix $\varepsilon > 0$, $0 < \delta < 1/\underline{\alpha}$ and write $z_n = n^{1-1/\underline{\alpha}+\delta}$. We observe that

$$\left| R_1^{\mathbf{f}}(j;n) \right| \leq \sum_{k=0}^{n-1} N_{ij}(k) \mathbb{1}_{\{u^{(i)}(k) \vdash u^{(i)}(n)\}}.$$

and

$$\#\{k \geq 0 : u^{(i)}(k) \vdash u^{(i)}(n)\} \leq \mathrm{Anc}_{\mathbf{f}}^{u^{(i)}(n)}(i) \leq \max_{0 \leq k \leq n} H_k^{\Pi^{(i)}(\mathbf{f})}.$$

Thus, according to our estimate for the height of GW forests in Lemma 4.2, we get that

$$\mathbf{P}^{\mathbf{x}} \left(\left| R_1^F(j;n) \right| > \varepsilon n^{1+\delta} \right) \quad \leq \quad \mathbf{P}^{\mathbf{x}} \left(\sum_{k=0}^{\lfloor z_n \rfloor} N_{ij}(k) > \varepsilon n^{1+\delta} \right) + \mathrm{oe}(n).$$

Moreover, for every $\beta \in (0, 1/2)$,

$$\begin{aligned} &\mathbf{P}^{\mathbf{x}} \left(\left| R_1^F(j;n) \right| > \varepsilon n^{1+\delta} \right) \\ &\quad \leq \mathbf{P}^{\mathbf{x}} \left(\left\{ \sum_{k=1}^{\lfloor z_n \rfloor} N_{ij}(k) > \varepsilon n^{1+\delta} \right\} \cap \left\{ \forall k \in \{0,1,\dots,\lfloor z_n \rfloor\} : N_{ij}(k) < (1-\beta)\varepsilon n^{1+\delta} \right\} \right) \\ &\qquad + \mathbf{P}^{\mathbf{x}} \left(\max_{1 \leq k \leq \lfloor z_n \rfloor} N_{ij}(k) > (1-\beta)\varepsilon n^{1+\delta} \right) + \mathrm{oe}(n). \end{aligned} \tag{4.10}$$

We recall that under $\mathbf{P}^{\mathbf{x}}$, the random variables $(N_{ij}(k), k \geq 0)$ are i.i.d. with law in the domain of attraction of a stable law of index $\underline{\alpha} \in (1, 2]$ by Proposition 4.2 (i). Then,

$$\mathbf{P}^{\mathbf{x}} \left(\max_{0 \leq k \leq \lfloor z_n \rfloor} N_{ij}(k) > (1-\beta)\varepsilon n^{1+\delta} \right) = 1 - \left(1 - \mathbf{P}^{\mathbf{x}} \left(N_{ij}(0) > (1-\beta)\varepsilon n^{1+\delta} \right) \right)^{\lfloor z_n \rfloor} = 0,$$

as $n \to \infty$. On the other hand, the first term in the right-hand side of (4.10) also tends to 0 as $n \to \infty$. To see this, note that the event in the first term may hold only if there are two distinct values of $k \in$

$\{0, 1, \ldots, \lfloor z_n \rfloor\}$ such that $N_{ij}(k) \geq \beta\varepsilon n / \lfloor z_n \rfloor$. We thus conclude that

$$\mathbf{P}^{\mathbf{x}}\left(\left|R_1^F(j;n)\right| > \varepsilon n^{1+\delta}\right) \to 0, \quad \text{as} \quad n \to \infty. \tag{4.11}$$

Following exactly the same argument, using the bound in Lemma 4.2 on the number of components of d-type GW forests and Proposition 4.2 (ii), we obtain that

$$\mathbf{P}^{\mathbf{x}}\left(\left|R_2^F(j;n)\right| > \varepsilon n^{1+\delta}\right) \to 0, \quad \text{as} \quad n \to \infty. \tag{4.12}$$

Finally, the estimate

$$\mathbf{P}^{\mathbf{x}}\left(\left|R_3^F(j;n)\right| > \varepsilon n^{1+\delta}\right) \to 0, \quad \text{as} \quad n \to \infty, \tag{4.13}$$

follows by the law of large numbers, since Proposition 4.2 (i) entails that the mean of $N_{ij}(0)$ is a_j/a_i.

Therefore, the estimates (4.11), (4.12) and (4.13), when combined with (4.9) imply the convergence (4.8). □

4.3 Proof of Theorem 4.1 and 4.2

In this section, we prove our main results.

Proof of Theorem 4.1. We observe that for $n \geq 0$ and any $s \geq 0$, we have

$$\left| H_{\lfloor ns \rfloor}^F - \frac{H^{\Pi^{(i)}(F)}_{\Lambda_i^F(\lfloor ns \rfloor)-1}}{a_i b_i} \right| \leq \left| H_{\lfloor ns \rfloor}^F - \frac{\mathrm{Anc}_F^{u(\lfloor ns \rfloor)}(i)}{a_i b_i} \right| + \frac{1}{a_i b_i}\left| H^{\Pi^{(i)}(F)}_{\Lambda_i^F(\lfloor ns \rfloor)-1} - \mathrm{Anc}_F^{u(\lfloor ns \rfloor)}(i) \right|.$$

By Proposition 4.3, under $\mathbf{P}^{\mathbf{x}}$, the first term on the right hand side tends to 0 in probability as $n \to \infty$, uniformly over compact subsets of $\mathbb{R}_+$. On the other hand, from equation (15) in [1], we get that

$$\left| H^{\Pi^{(i)}(F)}_{\Lambda_i^F(\lfloor ns \rfloor)-1} - \mathrm{Anc}_F^{u(\lfloor ns \rfloor)}(i) \right| \leq \left| H^{\Pi^{(i)}(F)}_{\Lambda_i^F(\lfloor ns \rfloor)-1} - H^{\Pi^{(i)}(F)}_{\Lambda_i^F(\lfloor ns \rfloor)} \right| + 1.$$

Recall that under $\mathbf{P}^{\mathbf{x}}$, $\Pi^{(i)}(F)$ is a critical non-degenerate monotype GW forest in the domain of attraction of a stable law of index $\underline{\alpha} \in (1, 2]$ by Proposition 4.1. Then, Theorem 2.3.2 in [62] implies that

$$\frac{1}{n^{1-1/\underline{\alpha}}} \max_{0 \leq k \leq n} \left| H_{k-1}^{\Pi^{(i)}(F)} - H_k^{\Pi^{(i)}(F)} \right| \underset{n\to\infty}{\to} 0,$$

in probability, under $\mathbf{P}^{\mathbf{x}}$, and it follows that

$$\left(\frac{1}{n^{1-1/\underline{\alpha}}} \left(H_{\lfloor ns \rfloor}^F - \frac{1}{a_i b_i} H^{\Pi^{(i)}(F)}_{\Lambda_i^F(\lfloor ns \rfloor)} \right), s \geq 0 \right) \underset{n\to\infty}{\to} 0 \tag{4.14}$$

in probability for the topology of uniform convergence over compact sets of $\mathbb{R}_+$. Finally, Proposition 4.4 and Theorem 2.3.2 in [62] imply that

$$\left(\frac{1}{n^{1-1/\underline{\alpha}}} H^{\Pi^{(i)}(F)}_{\Lambda_i^F(\lfloor ns \rfloor)}, s \geq 0\right) \xrightarrow[n\to\infty]{d} \left(\frac{a_i^{1/\underline{\alpha}} b_i}{\bar{c}} H_{a_i s}, s \geq 0\right).$$

Moreover, we deduce from the scaling property of the height process H that $(H_{a_i s}, s \geq 0) \stackrel{d}{=} (a_i^{1-1/\underline{\alpha}} H_s, s \geq 0)$; see, e.g., Section 3.1 in [62]. Therefore, the result in Theorem 4.1 follows now from (4.14). □

Let us now prove Theorem 4.2.

Proof of Theorem 4.2. For $n \geq 0$, $i \in [d]$ and any $s \geq 0$, we recall that $\Lambda_i^F(\lfloor ns \rfloor)$ denotes the number of type i individuals standing before the $(\lfloor ns \rfloor + 1)$-th individual in depth-first order which we called $u(\lfloor ns \rfloor)$. Since all the roots of the forest F have type i, we claim that

$$\Upsilon^{\Pi^{(i)}(F)}_{\Lambda_i^F(\lfloor ns \rfloor)} = \Upsilon_{\lfloor ns \rfloor}.$$

To see this, we observe that $u(\lfloor ns \rfloor)$ and the last vertex of type i before $u(\lfloor ns \rfloor)$ in depth-first order belong to the same tree component. Therefore, the label of the tree component of F containing $u(\lfloor ns \rfloor)$ is the same as the label of the tree component of $\Pi^{(i)}(F)$ containing the $\Lambda_i^F(\lfloor ns \rfloor)$-th vertex.

Let $W^{\Pi^{(i)}(F)} = (W_n^{\Pi^{(i)}(F)}, n \geq 1)$ be the Lukasiewicz path associated with monotype GW forest $\Pi^{(i)}(F)$ (see proof of Lemma 4.2 for the definition) which according to Proposition 4.1 has offspring distribution that belongs to the domain of attraction of a stable law of index $\underline{\alpha} \in (1, 2]$. We need the following property of Lukasiewicz path,

$$\inf_{0 \leq k \leq n} W_k^{\Pi^{(i)}(F)} = -\Upsilon_n^{\Pi^{(i)}(F)},$$

for $n \geq 1$; see for example [57]. The result now follows from Corollary 2.5.1 in [57] and similar arguments as at the end of proof of Theorem 4.1. □

4.4 Applications

4.4.1 Maximal height of multitype GW trees

In this section, we present a natural consequence of Theorems 4.1 and 4.2 which generalizes the result of Miermont [1] on the maximal height in the finite covariance case. For a tree $\mathbf{t} \in \mathbb{T}$, we let $\mathrm{ht}(\mathbf{t})$ be the maximal height of a vertex in $\mathbf{t}$. Recall that I_s is the infimum at time s of the strictly stable spectrally positive Lévy process $Y^{(\underline{\alpha})}$.

Corollary 4.1. *For $i \in [d]$, let T be a d-type GW tree distributed according to $\mathbf{P}^{(i)}$ whose offspring distribution satisfies the main assumptions. Then,*

$$\lim_{n\to\infty} n\mathbf{P}^{(i)}\left(\mathrm{ht}(T) \geq n\right) = b_i(\underline{\alpha} - 1)\left((\underline{\alpha} - 1)\bar{c}\right)^{\frac{\underline{\alpha}}{1-\underline{\alpha}}}.$$

Proof. The proof of this assertion is very similar of Corollary 1 in [1]. The only difference that we are now considering that the rescaled height process of multitype GW forest converges to height process associated with the strictly stable spectrally positive Lévy process $Y^{(\underline{\alpha})}$. Let F be a d-type GW forest distributed according to $\mathbf{P}^{(\mathbf{i})}$ whose offspring distribution satisfies the main assumptions. For $k \geq 1$, we denote by τ_k the first hitting time of k by $(\Upsilon^F_n, n \geq 0)$ and for $x \geq 0$, we write ϱ_x for the first hitting time of x by $-I = (-I_s, s \geq 0)$. From Theorem 4.1 and 4.2, we have that

$$\left(\frac{1}{n} H^F_{n^{\frac{\alpha}{\alpha-1}}s}, 0 \leq s \leq \tau_n\right) \xrightarrow[n\to\infty]{d} \left(\frac{1}{\bar{c}} H_s, 0 \leq s \leq \varrho_{b_i \bar{c}^{-1}}\right),$$

under $\mathbf{P}^{(\mathbf{i})}$. Let $(F_k, k \geq 1)$ be the tree components of the multitype GW forest F. Then, the above convergence implies that

$$\begin{aligned}
\lim_{n\to\infty} \mathbf{P}^{(\mathbf{i})} \left(\max_{1\leq k \leq n} \mathrm{ht}(F_k) < n\right) &= \mathbf{P}\left(H_s \leq \bar{c}, \text{ for all } 0 \leq s \leq \varrho_{b_i \bar{c}^{-1}}\right) \\
&= \exp\left(-\frac{b_i}{\bar{c}_i} N\left(\frac{1}{\bar{c}} \sup H \geq 1\right)\right) \\
&= \exp\left(-b_i(\underline{\alpha}-1)\left((\underline{\alpha}-1)\bar{c}\right)^{\frac{\alpha}{1-\alpha}}\right),
\end{aligned}$$

where N is the Itô excursion measure of $Y^{(\underline{\alpha})}$ above its infimum (see e.g. Chapter VIII.2 in [86] for details), and where we have used the Corollary 1.4.2 in [62] for the equality. Recall that under $\mathbf{P}^{(\mathbf{i})}$, the tree components $(F_k, k \geq 1)$ are independent multitype GW trees. Therefore, the identity

$$\mathbf{P}^{(\mathbf{i})} \left(\max_{1\leq k \leq n} \mathrm{ht}(F_k) < n\right) = \left(1 - \mathbf{P}^{(i)}\left(\mathrm{ht}(T) \geq n\right)\right)^n.$$

yields our claim. □

4.4.2 Alternating two-type GW tree

We consider a particular family of multitype GW trees known as alternating two-type GW trees, in which vertices of type 1 only give birth to vertices of type 2 and vice versa. More precisely, given two probability measures $\mu_2^{(1)}$ and $\mu_1^{(2)}$ on $\mathbb{Z}_+$, we consider a two-type GW tree where every vertex of type 1 (resp. type 2) has a number of type 2 (resp. type 1) children distributed according to $\mu_2^{(1)}$ (resp. $\mu_1^{(2)}$), all independent of each other. We denote by $\boldsymbol{\mu}_{\text{alt}}$ the offspring distribution on $\mathbb{Z}_+^2$ of this particular two-type GW tree. We let

$$m_{12} = \sum_{z\in\mathbb{Z}_+} z\mu_2^{(1)}(\{z\}) \quad \text{and} \quad m_{21} = \sum_{z\in\mathbb{Z}_+} z\mu_1^{(2)}(\{z\})$$

be the means of the measures $\mu_2^{(1)}$ and $\mu_1^{(2)}$, respectively. We make the assumption that $\mu_2^{(1)}(\{1\}) + \mu_1^{(2)}(\{1\}) < 2$ to exclude degenerate cases, and also exclude the trivial case $m_1 m_2 = 0$. We observe that the mean matrix associated with $\boldsymbol{\mu}_{\text{alt}}$ is irreducible and it admits $\rho = m_1 m_2$ as a unique positive eigenvalue. We then say that $\boldsymbol{\mu}_{\text{alt}}$ is sub-critical if $m_1 m_2 < 1$, critical if $m_1 m_2 = 1$ and supercritical if $m_1 m_2 > 1$. In the sequel, we assume that offspring distribution is also critical. We observe then that the normalized left and right 1-eigenvectors are given by

$$\mathbf{a} = (a_1, a_2) = \left(\frac{1}{1+m_1}, \frac{1}{1+m_2} \right), \quad \text{and} \quad \mathbf{b} = (b_1, b_2) = \left(\frac{1+m_1}{2}, \frac{1+m_2}{2} \right).$$

Following the notation of Section 4.1.3, we denote by $\mathbf{P}_{\text{alt}}^{(i)}$ the law of a two-type GW tree with offspring distribution $\boldsymbol{\mu}_{\text{alt}}$ and root type $i \in [2]$, i.e., it is the law of an alternating two-type GW tree with root type i. We make the next extra assumptions on the offspring distribution:

(H$'_1$) $\mu_2^{(1)}$ is a geometric distribution, i.e. there exists $p \in (0,1)$ such that

$$\mu_2^{(1)}(\{z\}) = (1-p)p^z, \qquad z \in \mathbb{Z}_+.$$

We observe that its Laplace exponent satisfies

$$\psi_1(s) = \frac{p}{1-p}s + \frac{1}{2}\frac{p}{(1-p)^2}s^2 + o(s^2), \qquad s \downarrow 0,$$

for $s \in \mathbb{R}_+$. In particular, $m_1 = p/(1-p)$.

(H$'_2$) $\mu_1^{(2)}$ is in the domain of attraction of a stable law of index $\alpha \in (1,2]$, that is, its Laplace exponent satisfies

$$\psi_2(s) = m_2 s + s^\alpha L(s) + o(s^\alpha), \qquad s \downarrow 0,$$

for $s \in \mathbb{R}_+$ and where $L : \mathbb{R}_+ \to \mathbb{R}_+$ is a slowly varying function at zero.

The following result is a conditioned version of Theorem 4.1 for this particular two-type GW tree. More precisely, we show that after a proper rescaling the height process of a critical alternating two-type GW tree whose offspring distribution satisfies (**H$'_1$**) and (**H$'_2$**) converges to the normalized excursion of the continuous-time height process associated with a strictly stable spectrally positive Lévy process with index α. We stress that the improvement of the convergence in Theorem 4.1 is because we are able to establish a conditioned version of Proposition 4.4 for this very particular GW tree. This allows us to adapt the proof of Theorem 2 in [1], in the case where only the geometric part of the offspring distribution does have small exponential moments.

Before providing a rigorous statement, we need to introduce some further notation. We consider a function $\bar{L} : \mathbb{R}_+ \to \mathbb{R}_+$ given by

$$\bar{L}(s) = \left(\frac{1}{2}\frac{p}{(1-p)^2} a_1 b_2^2 \mathbb{1}_{\{\alpha=2\}} + a_2 b_1^\alpha L(s) \right), \qquad \text{for } s \in \mathbb{R}_+, \tag{4.15}$$

which is a slowly varying function at zero. We write $\tilde{L} : \mathbb{R}_+ \to \mathbb{R}_+$ for a slowly varying function at infinity that satisfies

$$\lim_{s \to \infty} \left(\frac{1}{\tilde{L}(s)} \right)^\alpha \bar{L}\left(\frac{1}{s^{1/\alpha}\tilde{L}(s)} \right) = 1,$$

This function is known in the literature as the conjugate of $\bar{L}$. The existence of such a function is due to a result of de Bruijn; for a proof of this fact and more information about conjugate functions, see Section 1.5.7 in [87]. In what follows, we let $(B_n, n \geq 1)$ be a sequence positive integers such that $B_n = \tilde{L}(n)n^{1/\alpha}$.

Finally, recall from the beginning of Section that $H^{\mathbf{t}} = (H^{\mathbf{t}}_n, n \geq 0)$ denotes the height process of the tree $\mathbf{t} \in \mathbb{T}$.

Theorem 4.3. *Let T be an alternating two-type GW tree distributed according to $\mathbf{P}^{(1)}_{\text{alt}}$. Then for $j = 1, 2$, under the law $\mathbf{P}^{(1)}_{\text{alt}}(\cdot|\#T^{(j)} = n)$, the following convergence in distribution holds on $\mathbb{D}([0,1], \mathbb{R})$:*

$$\left(\frac{B_n}{n} H^T_{\lfloor \#Ts \rfloor}, 0 \leq s \leq 1 \right) \xrightarrow[n \to \infty]{d} \left(a_j^{1/\alpha - 1} H^{exc}_s, 0 \leq s \leq 1 \right),$$

where H^{exc} is the normalized excursion of the continuous-time height process process associated with a strictly stable spectrally positive Lévy process $Y^{(\alpha)} = (Y_s, s \geq 0)$ of index α and with Laplace exponent $\mathbb{E}(\exp(-\lambda Y_s)) = \exp(-s\lambda^\alpha)$, for $\lambda \in \mathbb{R}_+$.

In recent years, this special family of two-type GW trees has been the subject of many studies due to their remarkable relationship with the study of several important objects and models of growing relevance in modern probability such that random planar maps [70], percolation on random maps [71], non-crossing partitions [72], to mention just a few. On the other hand, up to our knowledge the result of Theorem 4.3 has not been proved before under our assumptions on the offspring distribution. Therefore, we believe that this may open the way to investigate new aspects related to the models mentioned before.

The proof of Theorem 4.3 relies on some intermediate results. We let T be a two-type GW tree with law $\mathbf{P}^{(1)}_{\text{alt}}$. We first characterize the law of the reduced forest $\Pi^{(j)}(T)$, for $j = 1, 2$.

Corollary 4.2. *For $j = 1, 2$, under the law $\mathbf{P}^{(1)}_{\text{alt}}$, the tree $\Pi^{(j)}(T)$ is a critical monotype GW forest with non-degenerate offspring distribution $\bar{\mu}_j$ in the domain of attraction of a stable law of index α, i.e., its Laplace exponent satisfies that*

$$\bar{\psi}_j(s) = s + \frac{1}{a_j} \left(\frac{s}{b_j} \right)^\alpha \bar{L}(s) + o(s^\alpha), \qquad s \downarrow 0.$$

for $s \in \mathbb{R}_+$ and where the function $\bar{L}$ is defined in (4.15).

Proof. The results follows from Lemma 4.1, after some simple computations. □

The next step in order to pass from unconditional statements to conditional ones is the following estimate for the number of vertices of some specific type in multitype GW trees.

Lemma 4.3. *Let T be a d-type GW tree distributed according to $\mathbf{P}^{(i)}$, for $i \in [d]$. Then, for every $j \in [d]$:*

(i) For some constant $C_{ij} > 0$, we have that

$$\mathbf{P}^{(i)} \left(\#T^{(j)} = n \right) = C_{ij} n^{-1-1/\underline{\alpha}} + o(n^{-1-1/\underline{\alpha}}), \qquad \text{as } n \to \infty,$$

where it is understood that the limit is taken along values for which the probability on the left-hand side is strictly positive.

(ii) The laws of the number of tree components of $\Pi^{(j)}(T)$, under $\mathbf{P}^{(i)}(\cdot|\#T^{(j)} = n)$, converge weakly as $n \to \infty$.

Proof. This very similar to Lemma 6 and Lemma 7 in [1] and the proof is carried out with mild modifications. □

Finally, the last ingredient is a conditioned version of Proposition 4.4 for the alternating two-type GW tree.

Proposition 4.5. *For $j = 1, 2$, under $\mathbf{P}^{(1)}_{\text{alt}}(\cdot|\#T^{(j)} = n)$, we have that*

$$\left(\frac{\Lambda^T_j(\lfloor \#Ts \rfloor)}{n}, 0 \leq s \leq 1\right) \underset{n\to\infty}{\to} (s, 0 \leq s \leq 1),$$

in probability.

Proof. We prove the statement only when $j = 1$. The case $j = 2$ follows by making occasional changes in the proof below, observing that

$$\Lambda^T_1(\#T) + \Lambda^T_2(\#T) = \#T^{(1)} + \#T^{(2)} = \#T.$$

We based our proof on a bijection $\mathcal{G}$ due to Janson and Stefánsson [88] which maps the alternating two-type GW tree to a standard monotype GW tree. More precisely, the tree $\mathcal{G}(T)$ has the same vertices as T, but edges are different and are defined as follows. For every type 1 vertex u we repeat the following operation: let u_0 be the parent of u (if $u \neq \varnothing$) and we list the children of u in lexicographical order $u_1 \prec u_2 \prec \cdots \prec u_k$. If $u \neq \varnothing$ draw the edge between u_0 and u_1 and then edges between u_1 and $u_2, \ldots, u_{k-1}$ and u_k and finally between u_k and u. If u is a type 1 vertex and a leaf this reduces to draw the edge between u_0 and u. One can check that $\mathcal{G}(T)$ defined by this procedure is a tree and rooted at the corner between the root of T and its first child. Roughly speaking, this mapping has the property that every vertex of type 1 is mapped to a leaf, and every type 2 vertex with $k \geq 0$ children is mapped to a vertex with $k + 1$ children (the interest reader is referred to Section 3 in [88], for details). Moreover, Janson and Stefánsson showed that under $\mathbf{P}^{(1)}_{\text{alt}}$, $\mathcal{G}(T)$ is a monotype GW tree with offspring distribution given by

$$\nu(\{0\}) = 1 - p, \quad \text{and} \quad \nu(\{z\}) = p\mu_2(\{z\}), \quad \text{for } z \in \mathbb{N}.$$

We notice that $\Lambda^T_1(\#T) = \#T^{(1)}$ is exactly the number of leaves of the monotype GW tree $\mathcal{G}(T)$. Then, Lemma 2.5 in [64] which is a law of large numbers for the number of leaves of monotype GW trees, implies that for every $\varepsilon > 0$,

$$\mathbf{P}^{(1)}_{\text{alt}}\left(\sup_{0\leq s\leq 1}\left|\frac{\Lambda^T_1(\lfloor \#Ts \rfloor)}{\#Ts} - (1-p)\right| > \varepsilon \middle| \#T \geq n\right) = \text{oe}(n).$$

We observe that the left 1-eigenvector $a_1 = 1 - p$. By Lemma 4.3, we deduce that

$$\mathbf{P}^{(1)}_{\text{alt}}\left(\sup_{0\le s\le 1}\left|\frac{\Lambda_1^T(\lfloor \#Ts\rfloor)}{\#Ts} - a_1\right| > \varepsilon \middle| \#T^{(1)} = n\right) = \text{oe}(n). \tag{4.16}$$

Then, if we admit for a while that

$$\mathbf{P}^{(1)}_{\text{alt}}\left(\left|\frac{\#T}{n} - \frac{1}{a_1}\right| > \varepsilon \middle| \#T^{(1)} = n\right) = \text{oe}(n). \tag{4.17}$$

We conclude the proof by combining the above estimate and (4.16).

Let us now turn to the proof of (4.17). First, we observe that for $0 < \varepsilon < a_1^{-1}$, we have that

$$\begin{aligned}\mathbf{P}^{(1)}_{\text{alt}}\left(\left|\frac{\#T}{n} - \frac{1}{a_1}\right| > \varepsilon, \#T^{(1)} = n\right) &= \mathbf{P}^{(1)}_{\text{alt}}\left(\#T > \left(\frac{1}{a_1} + \varepsilon\right)n, \#T^{(1)} = n\right)\\ &\quad + \mathbf{P}^{(1)}_{\text{alt}}\left(\#T < \left(\frac{1}{a_1} - \varepsilon\right)n, \#T^{(1)} = n\right).\end{aligned} \tag{4.18}$$

The idea is to show that the two term on the right-hand side are $\text{oe}(n)$. We start with the first term. We notice that

$$\mathbf{P}^{(1)}_{\text{alt}}\left(\#T > \left(\frac{1}{a_1} + \varepsilon\right)n, \#T^{(1)} = n\right) \le \sum_{k=n}^{\infty}\mathbf{P}^{(1)}_{\text{alt}}\left(\#T = k, \#T^{(1)} < \left(\frac{1}{a_1} + \varepsilon\right)^{-1} n\right)$$

By recalling that $\#T^{(1)}$ is the number of leaves of the monotype GW tree $\mathcal{G}(T)$, Lemma 2.7 (ii) in [64] implies that terms in the sum are $\text{oe}(n)$. This entails that the first term on the right-hand side of (4.18) is $\text{oe}(n)$. We now focus on the second term. We write

$$\mathbf{P}^{(1)}_{\text{alt}}\left(\#T > \left(\frac{1}{a_1} + \varepsilon\right)n, \#T^{(1)} = n\right) \le \sum_{k=n}^{\lfloor (a_1^{-1}-\varepsilon)n\rfloor}\mathbf{P}^{(1)}_{\text{alt}}\left(\#T = k, \#T^{(1)} > \left(\frac{1}{a_1} - \varepsilon\right)^{-1} n\right)$$

By using Proposition 1.6 in [64], we get that

$$\mathbf{P}^{(1)}_{\text{alt}}\left(\#T > \left(\frac{1}{a_1} + \varepsilon\right)n, \#T^{(1)} = n\right) \le \sum_{k=n}^{\lfloor (a_1^{-1}-\varepsilon)n\rfloor}\frac{1}{n}\mathbf{P}^{(1)}_{\text{alt}}\left(\frac{1}{r}\sum_{r=1}^{k}\mathbf{1}_{\{X_r=-1\}} > \left(\frac{1}{a_1} - \varepsilon\right)^{-1}\right),$$

where $(X_r, r \ge 1)$ is a sequence of i.i.d. random variables with common distribution $\nu(\{\cdot + 1\})$ on $\{-1\} \cup \mathbb{Z}_+$. Then, an application of Lemma 2.2 (i) in [64] shows that this is $\text{oe}(n)$. Therefore, we have proved that

$$\mathbf{P}^{(1)}_{\text{alt}}\left(\left|\frac{\#T}{n} - \frac{1}{a_1}\right| > \varepsilon, \#T^{(1)} = n\right) = \text{oe}(n). \tag{4.19}$$

Finally, an appeal to Lemma 4.3 (i) completes the proof of (4.17). □

We have now all the ingredients to give the proof of Theorem 4.3.

Proof of Theorem 4.3. Recall from Corollary 4.2 that $\Pi^{(j)}(T)$ under $\mathbf{P}_{\text{alt}}^{(1)}$ is a non-degenerate, critical GW forest with offspring distribution $\bar{\mu}_j$ in the domain of attraction of a stable law of index $\alpha \in (1,2]$. Thus, by first conditioning on the number of tree components, we obtain using Lemma 4.3 (ii) and Theorem 3.1 [57] that under $\mathbf{P}_{\text{alt}}^{(1)}(\cdot | \#T^{(j)} = n)$,

$$\left(\frac{B_n}{n} H_{\lfloor ns \rfloor}^{\Pi^{(j)}(T)}, 0 \le s \le 1\right) \xrightarrow[n\to\infty]{d} \left(a_j^{1/\alpha} b_j H_s^{\text{exc}}, 0 \le s \le 1\right),$$

where the convergence is in distribution on $\mathbb{D}([0,1],\mathbb{R})$. To see this, we observe that conditional on the number of tree components to be r, the GW forest $\Pi^{(j)}(T)$ is composed of r independent GW trees with the same offspring distribution $\bar{\mu}_j$. On the other hand, conditioning the sum of their size to be n, only one of these trees has size of order n, while the other $r-1$ trees have total size $o(n)$ with high probability. This implies that the latter do not contribute to the limit. We refer to Theorem 5.4 in [89] for details. Then, from Proposition 4.5, we obtain that under $\mathbf{P}_{\text{alt}}^{(1)}(\cdot | \#T^{(j)} = n)$,

$$\left(\frac{B_n}{n} H_{\Lambda_j^T(\lfloor \#Ts \rfloor)}^{\Pi^{(j)}(T)}, 0 \le s \le 1\right) \xrightarrow[n\to\infty]{d} \left(a_j^{1/\alpha} b_j H_s^{\text{exc}}, 0 \le s \le 1\right), \tag{4.20}$$

in distribution.

On the other hand, recall from the proof of Theorem 4.1 that for $n \ge 0$ and any $s \ge 0$, we have

$$\left| H_{\lfloor \#Ts \rfloor}^T - \frac{H_{\Lambda_j^T(\lfloor \#Ts \rfloor)}^{\Pi^{(j)}(T)}}{a_j b_j} \right| \le \left| H_{\lfloor \#Ts \rfloor}^T - \frac{\text{Anc}_T^{u(\lfloor \#Ts \rfloor)}(j)}{a_j b_j} \right| + R_n(s), \tag{4.21}$$

where

$$|R_n(s)| \le \frac{1}{a_j b_j} \left(2 \max_{0 \le k \le n} \left| H_{k-1}^{\Pi^{(j)}(T)} - H_k^{\Pi^{(j)}(T)} \right| + 1 \right).$$

Therefore, it must be clear that our claim follows from the convergence (4.20) by providing that the two terms on the right-hand side of (4.21) are $o(n/B_n)$ in probability, uniformly in $s \in [0,1]$.

In this direction, we observe from (4.19) that $\mathbf{P}_{\text{alt}}^{(1)}(\#T > \delta n | \#T^{(j)} = n) = \text{oe}(n)$ for any $\delta > a_j^{-1}$. Combining this with Proposition 4.3, we have for $0 < \gamma < \frac{1}{2}(1 - 1/\alpha)$ and some $C > 0$ that

$$\begin{aligned} &\mathbf{P}_{\text{alt}}^{(1)}\left(\frac{B_n}{n} \max_{0 \le k \le \#T} \left| H_k^T - \frac{\text{Anc}_T^{u(k)}(j)}{a_j b_j} \right| \ge n^{-\frac{1}{2}(1-1/\alpha)+\gamma} \middle| \#T^{(j)} = n \right) \\ &\le C n^{1+1/\alpha} \mathbf{P}_{\text{alt}}^{(\mathbf{1})}\left(\frac{B_n}{n} \max_{0 \le k \le \lfloor \delta n \rfloor} \left| H_k^T - \frac{\text{Anc}_T^{u(k)}(j)}{a_j b_j} \right| \ge n^{-\frac{1}{2}(1-1/\alpha)+\gamma} \right) + \text{oe}(n) = \text{oe}(n), \end{aligned}$$

where $\mathbf{P}_{\text{alt}}^{(\mathbf{1})}$ is the law of alternating two-type GW forest with all its root having type 1. This shows that first term on the right-hand side of (4.21) is $o(n/B_n)$ in probability, uniformly in $s \in [0,1]$.

Finally, let Υ^j be the number of tree components of $\Pi^{(j)}(T)$. Then the law of $\Pi^{(j)}(T)$ under the measure $\mathbf{P}^{(1)}_{\text{alt}}(\cdot|\Upsilon^j = r)$ is that of a monotype GW forest with r tree components. Using Theorem 5.4 in [89], one concludes that for $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbf{P}^{(1)}_{\text{alt}}\left(\sup_{0\le s\le 1} \frac{B_n}{n}|R_n(s)| \ge \varepsilon \middle| \#T^{(j)} = n, \Upsilon^j = r\right) = 0.$$

By Lemma 4.3 (ii), we know that the laws of Υ^j under $\mathbf{P}^{(1)}_{\text{alt}}(\cdot|\#T^{(j)} = n)$ are tight as n varies. Thus, we deduce that the second term on the right-hand side of (4.21) is also $o(n/B_n)$ in probability, uniformly in $s \in [0,1]$. $\square$

APPENDIX A

The continuous-time height process

We define the paths which are the analogs in continuous-time of the Lukasiewicz path and the discrete height process introduced in Section 1.4.1. We recall some definitions, properties and constructions with no proof and refer the interesting reader to Bertoin [86] and Duquesne [62] for more details. In this direction, let us introduce first some required notation. Let $I \subset \mathbb{R}$ be an interval. We denote by $\mathbb{C}(I, \mathbb{R})$ the space of real-valued continuous functions on I equipped with the uniform distance on all the compact subsets of I, which makes it a Polish space. We also denote by $\mathbb{D}(I, \mathbb{R})$ the space of càdlàd functions on I endowed with the Skorokhod topology, making it a Polish space (see Chapter 3 in [46] and Chapter IV in [90] for definitions and usual properties of the Skorokhod topology).

Fix $\alpha \in (1, 2]$ and consider a random process $Y^{(\alpha)} = (Y_t, t \geq 0)$ with paths in the set $\mathbb{D}([0, \infty), \mathbb{R})$, which has independent and stationary increments, no negative jump and such that

$$\mathbb{E}[\exp(-\lambda Y_t)] = \exp(t\lambda^\alpha), \qquad \text{for } \lambda > 0.$$

Such a process is called a strictly stable spectrally positive Lévy process of index α. An important feature of Y^α is the scaling property: for every $c > 0$,

$$(c^{-1/\alpha} Y_{ct}, t \geq 0) = (Y_t, t \geq 0) \quad \text{ in distribution.}$$

In particular, for $\alpha = 2$ the process $Y^{(2)}$ is $\sqrt{2}$ times the standard Brownian motion on the line. For $0 \leq s \leq t$, we set

$$I_{s,t} = \inf_{s \leq r \leq t} Y_r \quad \text{ and } \quad I_t = \inf_{0 \leq r \leq t} Y_r.$$

We write $H^{(\alpha)} = (H_t, t \geq 0)$ for the continuous-time height process associated with $Y^{(\alpha)}$. In the Brownian case, the height process is $H_t = X_t - I_t$ and obviously has continuous paths. Otherwise, we define the height process by

$$H_t = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^t \mathbb{1}_{\{Y_s < I_{s,t} + \varepsilon\}} \mathrm{d}s,$$

where the limit exists in probability. The process $H^{(\alpha)}$ admits a continuous modification by Theorem 4.7 in [56], and from now on we consider only this modification. Let us give a briefly explanation for this definition; see Chapter 1 in [62] for further details. For $t > 0$, we let $\hat{Y}^{(t)} = (\hat{Y}_s^{(t)}, 0 \leq s \leq t)$ be the time-reversed process of $Y^{(\alpha)}$ defined by $\hat{Y}_s^{(t)} = Y_t - Y_{(t-s)-}$. We set $\hat{S}_s^{(t)} = \sup_{0 \leq r \leq s} \hat{Y}_r^{(t)}$. It is possible to prove that H_t has the same law as the local time at time t (suitably normalized) of the process $\hat{S}^{(t)} - \hat{Y}^{(t)}$. Then H_t corresponds intuitively to the "measure" of the set $\{0 \leq s \leq t : Y_s = I_{s,t}\}$, by analogy with the discrete case (1.6). The role of the Lukasiewicz path is played by the stable Lévy process $Y^{(\alpha)}$. Next, one can deduce from the scaling property of $Y^{(\alpha)}$ that $H^{(\alpha)}$ satisfies

$$(c^{-1+1/\alpha} H_{ct}, t \geq 0) = (H_t, t \geq 0) \quad \text{in distribution.}$$

In addition, the excursions of $H^{(\alpha)}$ above 0 coincide with excursions of $Y^{(\alpha)} - I$ above 0, where $I = (I_s, s \geq 0)$.

The normalized excursions $Y_\alpha^{\text{exc}} = (Y_t^{\text{exc}}, t \geq 0)$ and $H_\alpha^{\text{exc}} = (H_t^{\text{exc}}, t \geq 0)$ are defined from $Y^{(\alpha)}$ and $H^{(\alpha)}$ as follows. We consider the times $\underline{g}_1 = \sup\{s \geq 1 : Y_s = I_s\}$ and $\zeta_1 = \inf\{s > 1 : Y_s = I_s\} - \underline{g}_1$. For $0 \leq t \leq 1$, we then define:

$$(Y_t^{\text{exc}}, H_t^{\text{exc}}) = \left(\zeta_1^{-\frac{1}{\alpha}} \left(Y_{\underline{g}_1 + \zeta_1 t} - Y_{\underline{g}_1} \right), \zeta_1^{\frac{1}{\alpha}-1} H_{\underline{g}_1 + \zeta_1 t} \right).$$

The process H_α^{exc} is called the normalized excursion of the height process $H^{(\alpha)}$. Moreover, almost surely we have that $H_0^{\text{exc}} = H_1^{\text{exc}} = 0$ and $H_t^{\text{exc}} > 0$ for all $0 < t < 1$. On the other hand, the process H_α^{exc} takes values in the Polish space $\mathbb{C}([0,1], \mathbb{R})$, and X_α^{exc} takes value in the set $\mathbb{D}([0,1], \mathbb{R})$. We stress that in the Brownian case $H_2^{\text{exc}} = \sqrt{2} \cdot \mathbb{e}$, where $\mathbb{e}$ is the normalized Brownian excursion.

Bibliography

[1] G. Miermont. Invariance principles for spatial multitype Galton-Watson trees. *Ann. Inst. Henri Poincaré Probab. Stat.* 44.6 (2008), pp. 1128–1161.

[2] G. Berzunza. Yule processes with rare mutation and their applications to percolation on b-ary trees. *Electron. J. Probab.* 20 (2015), no. 43, 1–23.

[3] G. Berzunza. The cut-tree of large trees with small heights. *Accepted at Random Structures Algorithms,* (2016).

[4] G. Berzunza. On scaling limits of multitype Galton-Watson trees with possibly infinite variance. *Preprint available at arXiv:1605.04810* (2016).

[5] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.* 53 (1947), pp. 292–294.

[6] P. Erdős. Graph theory and probability. *Canad. J. Math.* 11 (1959), pp. 34–38.

[7] P. Erdős. Graph theory and probability. II. *Canad. J. Math.* 13 (1961), pp. 346–352.

[8] M. Drmota. Random trees. Springer WienNewYork, Vienna, 2009.

[9] J. W. Moon. The distance between nodes in recursive trees. In: *Combinatorics (Proc. British Combinatorial Conf., Univ. Coll. Wales, Aberystwyth, 1973)*. Cambridge Univ. Press, London, 1974, 125–132. London Math. Soc. Lecture Note Ser., No. 13.

[10] P. Flajolet and J.-M. Steyaert. A complexity calculus for recursive tree algorithms. *Math. Systems Theory* 19.4 (1987), pp. 301–331.

[11] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science* 286.5439 (1999), pp. 509–512.

[12] J. Szymański. On a nonuniform random recursive tree. In: *Random graphs '85 (Poznań, 1985)*. Vol. 144. North-Holland Math. Stud. North-Holland, Amsterdam, 1987, pp. 297–306.

[13] H. M. Mahmoud. Evolution of random search trees. Wiley-Interscience Series in Discrete Mathematics and Optimization. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1992.

[14] J. Neveu. Arbres et processus de Galton-Watson. *Ann. Inst. H. Poincaré Probab. Statist.* 22.2 (1986), pp. 199–207.

[15] R. Otter. The multiplicative process. *Ann. Math. Statistics* 20 (1949), pp. 206–224.

[16] R. Abraham, J.-F. Delmas, and H. Guo. Critical multi-type Galton-Watson trees conditioned to be large. *Preprint available at arXiv:1511.01721* (2015).

[17] K. B. Athreya and P. E. Ney. Branching processes. Reprint of the 1972 original [Springer, New York; MR0373040]. Dover Publications, Inc., Mineola, NY, 2004.

[18] J. Bertoin. Almost giant clusters for percolation on large trees with logarithmic heights. *J. Appl. Probab.* 50.3 (2013), pp. 603–611.

[19] D. Aldous. The continuum random tree. I. *Ann. Probab.* 19.1 (1991), pp. 1–28.

[20] R. P. Dobrow. On the distribution of distances in recursive trees. *J. Appl. Probab.* 33.3 (1996), pp. 749–757.

[21] J. Schweinsberg. Dynamics of the evolving Bolthausen-Sznitman coalescent. *Electron. J. Probab.* 17 (2012), no. 91, 50.

[22] J. Bertoin. On the non-Gaussian fluctuations of the giant cluster for percolation on random recursive trees. *Electron. J. Probab.* 19 (2014), no. 24, 15.

[23] M. Möhle. Convergence results for compound Poisson distributions and applications to the standard Luria-Delbrück distribution. *J. Appl. Probab.* 42.3 (2005), pp. 620–631.

[24] A. Meir and J. W. Moon. Cutting down random trees. *J. Austral. Math. Soc.* 11 (1970), pp. 313–324.

[25] M. Drmota et al. A limiting distribution for the number of cuts needed to isolate the root of a random recursive tree. *Random Structures Algorithms* 34.3 (2009), pp. 319–336.

[26] A. Iksanov and M. Möhle. A probabilistic proof of a weak limit law for the number of cuts needed to isolate the root of a random recursive tree. *Electron. Comm. Probab.* 12 (2007), pp. 28–35.

[27] C. Holmgren. Random records and cuttings in binary search trees. *Combin. Probab. Comput.* 19.3 (2010), pp. 391–424.

[28] C. Holmgren. A weakly 1-stable distribution for the number of random records and cuttings in split trees. *Adv. in Appl. Probab.* 43.1 (2011), pp. 151–177.

[29] L. Devroye. Universal limit laws for depths in random trees. *SIAM J. Comput.* 28 (1999), pp. 409–432.

[30] C. Goldschmidt and J. B. Martin. Random recursive trees and the Bolthausen-Sznitman coalescent. *Electron. J. Probab.* 10 (2005), no. 21, 718–745.

[31] B. Chauvin et al. Martingales and profile of binary search trees. *Electron. J. Probab.* 10 (2005), no. 12, 420–435.

[32] J. Bertoin and G. Uribe Bravo. Supercritical percolation on large scale-free random trees. *Ann. Appl. Probab.* 25.1 (2015), pp. 81–103.

[33] A. Meir and J. Moon. Cutting down recursive trees. *Mathematical Biosciences* 21 (1974), pp. 173–181.

[34] A. Panholzer. Cutting down very simple trees. *Quaest. Math.* 29.2 (2006), pp. 211–227.

[35] S. Janson. Random cutting and records in deterministic and random trees. *Random Structures Algorithms* 29 (2006), pp. 139–179.

[36] S. Janson. Random records and cuttings in complete binary trees. In: *Mathematics and computer science. III*. Trends Math. Birkhäuser, Basel, 2004, pp. 241–253.

[37] J. Bertoin. Fires on trees. *Ann. Inst. Henri Poincaré Probab. Stat.* 48 (2012), pp. 909–921.

[38] J. Bertoin and G. Miermont. The cut-tree of large Galton-Watson trees and the Brownian CRT. *Ann. Appl. Probab.* 23 (2013), pp. 1469–1493.

[39] J. Bertoin. The cut-tree of large recursive trees. *Ann. Inst. Henri Poincaré Probab. Stat.* 51.2 (2015), pp. 478–488.

[40] D. Dieuleveut. The vertex-cut-tree of Galton-Watson trees converging to a stable tree. *Ann. Appl. Probab.* 25.4 (2015), pp. 2215–2262.

[41] N. Broutin and M. Wang. Cutting down p-trees and inhomogeneous continuum random trees. *Preprint available at arXiv:1408.0144* (2014).

[42] L. Addario-Berry, D. Dieuleveut, and C. Goldschmidt. Inverting the cut-tree transform, *Preprint available at arXiv:1606.04825* (2016).

[43] A. Dress, V. Moulton, and W. Terhalle. T-theory: an overview. *European J. Combin.* 17.2-3 (1996). Discrete metric spaces (Bielefeld, 1994), pp. 161–175.

[44] S. N. Evans. Probability and real trees. Vol. 1920. Lecture Notes in Mathematics. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005. Springer, Berlin, 2008.

[45] T. Duquesne and J.-F. Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields* 131.4 (2005), pp. 553–603.

[46] P. Billingsley. Convergence of probability measures. Second ed. Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999.

[47] S. N. Evans and A. Winter. Subtree prune and regraft: a reversible real tree-valued Markov process. *Ann. Probab.* 34.3 (2006), pp. 918–961.

[48] A. Greven, P. Pfaffelhuber, and A. Winter. Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees). *Probab. Theory Related Fields* 145.1-2 (2009), pp. 285–322.

[49] W. Löhr. Equivalence of Gromov-Prohorov- and Gromov's $\square_\lambda$-metric on the space of metric measure spaces. *Electron. Commun. Probab.* 18 (2013), no. 17, 10.

[50] M. Gromov. Metric structures for Riemannian and non-Riemannian spaces. Vol. 152. Progress in Mathematics. Birkhäuser Boston, MA MR 1699320, 1999.

[51] B. Haas and G. Miermont. Scaling limits of Markov branching trees with applications to Galton-Watson and random unordered trees. *Ann. Probab.* 40 (2012), pp. 2589–2666.

[52] J. Bertoin. Sizes of the largest clusters for supercritical percolation on random recursive trees. *Random Structures Algorithms* 44 (2014), pp. 29–44.

[53] M. Kuba and A. Panholzer. Multiple isolation of nodes in recursive trees. *Online J. Anal. Comb.* (2014), p. 26.

[54] E. Baur and J. Bertoin. Cutting Edges at Random in Large Recursive Trees. In: *Stochastic Analysis and Applications 2014*. Ed. by D. Crisan, B. Hambly, and T. Zariphopoulou. Vol. 100. Springer Proceedings in Mathematics & Statistics. Springer International Publishing, 2014, 51–76.

[55] T. E. Harris. The theory of branching processes. Dover Phoenix Editions. Corrected reprint of the 1963 original [Springer, Berlin; MR0163361 (29 #664)]. Dover Publications, Inc., Mineola, NY, 2002.

[56] J.-F. Le Gall and Y. Le Jan. Branching processes in Lévy processes: the exploration process. *Ann. Probab.* 26.1 (1998), pp. 213–252.

[57] T. Duquesne. A limit theorem for the contour process of conditioned Galton-Watson trees. *Ann. Probab.* 31.2 (2003), pp. 996–1027.

[58] G. Miermont. Tessellations of random maps of arbitrary genus. *Ann. Sci. Éc. Norm. Supér. (4)* 42.5 (2009), pp. 725–781.

[59] S. N. Evans, J. Pitman, and A. Winter. Rayleigh processes, real trees, and root growth with re-grafting. *Probab. Theory Related Fields* 134.1 (2006), pp. 81–126.

[60] D. Revuz and M. Yor. Continuous martingales and Brownian motion. Third. Vol. 293. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1999.

[61] D. Aldous. The continuum random tree. III. *Ann. Probab.* 21.1 (1993), pp. 248–289.

[62] T. Duquesne and J.-F. Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque* 281 (2002), pp. vi+147.

[63] J. Pitman. Combinatorial stochastic processes. Vol. 1875. Lecture Notes in Mathematics. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard. Springer-Verlag, Berlin, 2006.

[64] I. Kortchemski. Invariance principles for Galton-Watson trees conditioned on the number of leaves. *Stochastic Process. Appl.* 122.9 (2012), pp. 3126–3172.

[65] L. de Raphelis. Scaling limit of multitype Galton-Watson trees with infinitely many types. *To appear in Ann. Inst. Henri Poincaré Probab. Stat.* (2016).

[66] K.-i. Sato. Lévy processes and infinitely divisible distributions. Vol. 68. Cambridge Studies in Advanced Mathematics. Translated from the 1990 Japanese original, Revised edition of the 1999 English translation. Cambridge University Press, Cambridge, 2013.

[67] G. Samorodnitsky and M. S. Taqqu. Stable non-Gaussian random processes. Stochastic Modeling. Stochastic models with infinite variance. Chapman & Hall, New York, 1994.

[68] I. Kortchemski. A simple proof of Duquesne's theorem on contour processes of conditioned Galton-Watson trees. In: *Séminaire de Probabilités XLV*. Vol. 2078. Lecture Notes in Math. Springer, Cham, 2013, pp. 537–558.

[69] J.-F. Marckert and G. Miermont. Invariance principles for random bipartite planar maps. *Ann. Probab.* 35.5 (2007), pp. 1642–1705.

[70] J.-F. Le Gall and G. Miermont. Scaling limits of random planar maps with large faces. *Ann. Probab.* 39.1 (2011), pp. 1–69.

[71] N. Curien and I. Kortchemski. Percolation on random triangulations and stable looptrees. *Probab. Theory Related Fields* 163.1-2 (2015), pp. 303–337.

[72] I. Kortchemski and C. Marzouk. Simply generated non-crossing partitions. *To appear in Combinatorics, Probability and Computing.* (2016).

[73] M. Javanian and M. Q. Vahidi-Asl. Depth of nodes in random recursive k-ary trees. *Inform. Process. Lett.* 98.3 (2006), pp. 115–118.

[74] E. Parzen. Stochastic processes. Reprint of the 1962 original. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.

[75] A. de La Fortelle. Yule process sample path asymptotics. *Electron. Comm. Probab.* 11 (2006), pp. 193–199.

[76] E. Baur. Percolation on random recursive trees. To appear in *Random Structures Algorithms* (2016).

[77] A. Greven, P. Pfaffelhuber, and A. Winter. Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees). *Probab. Theory Related Fields* 145.1-2 (2009), pp. 285–322.

[78] L. Devroye. Applications of the theory of records in the study of random trees. *Acta Inform.* 26.1-2 (1988), pp. 123–130.

[79] Z. Katona. Width of a scale-free tree. *J. Appl. Probab.* 42.3 (2005), pp. 839–850.

[80] Z. Katona. Levels of a scale-free tree. *Random Structures Algorithms* 29.2 (2006), pp. 194–207.

[81] T. F. Móri. The maximum degree of the Barabási-Albert random tree. *Combin. Probab. Comput.* 14.3 (2005), pp. 339–348.

[82] J.-F. Le Gall. Random trees and applications. *Probab. Surv.* 2 (2005), pp. 245–311.

[83] P. Jagers. General branching processes as Markov fields. *Stochastic Process. Appl.* 32.2 (1989), pp. 183–212.

[84] W. Feller. An introduction to probability theory and its applications. Vol. II. Second edition. John Wiley & Sons, Inc., New York-London-Sydney, 1971.

[85] R. A. Doney. On the exact asymptotic behaviour of the distribution of ladder epochs. *Stochastic Process. Appl.* 12.2 (1982), pp. 203–214.

[86] J. Bertoin. Lévy processes. Vol. 121. Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1996, pp. x+265.

[87] N. H. Bingham, C. M. Goldie, and J. L. Teugels. Regular variation. Vol. 27. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1989.

[88] S. Janson and S. Ö. Stefánsson. Scaling limits of random planar maps with a unique large face. *Ann. Probab.* 43.3 (2015), pp. 1045–1081.

[89] I. Kortchemski and C. Marzouk. Triangulating stable laminations. *Electron. J. Probab.* 21 (2016), 31 pp.

[90] J. Jacod and A. N. Shiryaev. Limit theorems for stochastic processes. Vol. 288. Grundlehren der Mathematischen Wissenschaften. Springer-Verlag, Berlin, 1987.